

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 2, 2006, 14:00:24 ; Search time 163 Seconds
(without alignments)
64.323 Million cell updates/sec

Perfect score: 20
Sequence: 1 gagtaggaaggatagaacg 20

Title: US-09-675-650-4

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4172979 seqs, 262114271 residues

Total number of hits satisfying chosen parameters: 8345958

Minimum DB Seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

RESULT 1
US-10-773-440A-7/c
; Sequence 7, Application US/10773440A
; Publication No. US20050282170A1
; GENERAL INFORMATION:
; APPLICANT: Fradet, Yves
; APPLICANT: Chyfre, Camille
; APPLICANT: Piche, Lyson
; APPLICANT: Garon, Genevieve
; TITLE OF INVENTION: Method to Detect Prostate Cancer in a Sample
; FILE REFERENCE: 1619_018001
; CURRENT APPLICATION NUMBER: US/10/773,440A
; CURRENT FILING DATE: 2004-02-09
; PRIOR APPLICATION NUMBER: 60/1445,436
; PRIOR FILING DATE: 2003-02-07
; NUMBER OF SEQ ID NOS: 13
; SEQ ID NO: 7
; LENGTH: 506
; ORGANISM: Homo sapiens
US-10-773-440A-7

ALIGNMENTS

Result No.	Score	Query Match	Length	DB ID	Description
c 1	20	100.0	506	6 US-10-773-440A-7	Sequence 7, Appli
c 2	16.8	84.0	789	6 US-10-467-657-4593	Sequence 4593, Appli
c 3	16.8	84.0	3466	6 US-10-750-105-51092	Sequence 51092, Appli
c 4	16.4	82.0	1165	6 US-10-750-105-64441	Sequence 64341, Appli
c 5	16.4	82.0	1931	6 US-10-750-105-28759	Sequence 28759, Appli
c 6	15.8	79.0	2376	6 US-10-750-105-42164	Sequence 42164, Appli
c 7	15.8	79.0	27032	6 US-10-995-561-13468	Sequence 13468, Appli
c 8	15.8	79.0	260209	6 US-10-933-095-23	Sequence 23, Appli
c 9	15.8	79.0	611587	7 US-11-117-187-209	Sequence 209, Appli
c 10	15.4	77.0	1861	6 US-10-750-105-33082	Sequence 33082, Appli
c 11	15.4	77.0	3065	6 US-10-750-105-40303	Sequence 40303, Appli
c 12	15.4	77.0	645179	6 US-10-995-561-13493	Sequence 13493, Appli
c 13	15.2	76.0	201	6 US-10-995-561-14464	Sequence 14464, Appli
c 14	15.2	76.0	687	6 US-10-750-105-48039	Sequence 48039, Appli
c 15	15.2	76.0	1406	6 US-10-750-105-29945	Sequence 29945, Appli
c 16	15.2	76.0	1904	6 US-10-750-105-26121	Sequence 26121, Appli
c 17	15.2	76.0	2465	7 US-11-108-172-1049	Sequence 1049, Appli
c 18	15.2	76.0	40387	6 US-10-995-561-13503	Sequence 13503, Appli
c 19	15.2	76.0	67088	6 US-10-995-561-13465	Sequence 13465, Appli
c 20	15.2	76.0	96128	6 US-10-995-561-13197	Sequence 13197, Appli
c 21	15.2	76.0	120697	7 US-11-121-086-48	Sequence 85968, Appli
c 22	15	75.0	19	8 US-11-101-244-85687	Sequence 859687, Appli
c 23	15	75.0	19	9 US-11-083-784-29548	Sequence 29548, Appli

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	20	100.0	506	6 US-10-773-440A-7	Sequence 7, Appli
c 2	16.8	84.0	789	6 US-10-467-657-4593	Sequence 4593, Appli
c 3	16.8	84.0	3466	6 US-10-750-105-51092	Sequence 51092, Appli
c 4	16.4	82.0	1165	6 US-10-750-105-64441	Sequence 64341, Appli
c 5	16.4	82.0	1931	6 US-10-750-105-28759	Sequence 28759, Appli
c 6	15.8	79.0	2376	6 US-10-750-105-42164	Sequence 42164, Appli
c 7	15.8	79.0	27032	6 US-10-995-561-13468	Sequence 13468, Appli
c 8	15.8	79.0	260209	6 US-10-933-095-23	Sequence 23, Appli
c 9	15.8	79.0	611587	7 US-11-117-187-209	Sequence 209, Appli
c 10	15.4	77.0	1861	6 US-10-750-105-33082	Sequence 33082, Appli
c 11	15.4	77.0	3065	6 US-10-750-105-40303	Sequence 40303, Appli
c 12	15.4	77.0	645179	6 US-10-995-561-13493	Sequence 13493, Appli
c 13	15.2	76.0	201	6 US-10-995-561-14464	Sequence 14464, Appli
c 14	15.2	76.0	687	6 US-10-750-105-48039	Sequence 48039, Appli
c 15	15.2	76.0	1406	6 US-10-750-105-29945	Sequence 29945, Appli
c 16	15.2	76.0	1904	6 US-10-750-105-26121	Sequence 26121, Appli
c 17	15.2	76.0	2465	7 US-11-108-172-1049	Sequence 1049, Appli
c 18	15.2	76.0	40387	6 US-10-995-561-13503	Sequence 13503, Appli
c 19	15.2	76.0	67088	6 US-10-995-561-13465	Sequence 13465, Appli
c 20	15.2	76.0	96128	6 US-10-995-561-13197	Sequence 13197, Appli
c 21	15.2	76.0	120697	7 US-11-121-086-48	Sequence 85968, Appli
c 22	15	75.0	19	8 US-11-101-244-85687	Sequence 859687, Appli
c 23	15	75.0	19	9 US-11-083-784-29548	Sequence 29548, Appli

RESULT 2
US-10-467-657-4593/c
; Sequence 4593, Application US/10467657
; GENERAL INFORMATION:
; APPLICANT: CHIRONE SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZAI Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11

; PRIOR APPLICATION NUMBER: GB-0103424.8
 ; PRIORITY FILING DATE: 2001-02-12
 ; NUMBER OF SEQ ID NOS: 9218
 ; SOFTWARE: SeqWin99, version 1.04
 ; SEQ ID NO: 4593
 ; LENGTH: 789
 ; TYPE: DNA
 ; ORGANISM: Neisseria gonorrhoeae
 US-10-467-657-4593

Query Match Score 84.0%; DB 6; Length 3466;
 Best Local Similarity 90.0%; Pred. No. 50;
 Matches 18; Conservative 0; Mismatches 2;
 Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATCAAACG 20
 Db 148 GAGTAGGAATGAAACAACG 129

RESULT 3
 US-10-750-185-51092

; Sequence 51092, Application US/10750185
 ; Publication No. US20050260603A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MMI GENOMICS, INC.
 ; ATTORNEY: DENISE, Sue K.
 ; ATTORNEY: KERR, Richard
 ; ATTORNEY: ROSENFIELD, David
 ; ATTORNEY: HOLM, Tom
 ; ATTORNEY: BATES, Stephen
 ; ATTORNEY: FANTIN, Dennis
 ; FILE REFERENCE: MMII1100-2
 ; CURRENT APPLICATION NUMBER: US/10/750,185
 ; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
 ; FILE REFERENCE: MMII1100-2
 ; CURRENT FILING DATE: 2003-12-31
 ; PRIOR APPLICATION NUMBER: US 60/437,482
 ; PRIOR FILING DATE: 2002-12-31
 ; SOFTWARE: PatentIN version 3.1
 ; SEQ ID NO: 51092
 ; SEQ ID NO: 51092
 ; LENGTH: 3466
 ; TYPE: DNA
 ; ORGANISM: Bovine 19866880511595

Query Match Score 84.0%; DB 6; Length 3466;
 Best Local Similarity 90.0%; Pred. No. 50;
 Matches 18; Conservative 0; Mismatches 2;
 Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATCAAACG 20
 Db 2449 GAGTAGGAAGGATCAAACG 2466

RESULT 4
 US-10-750-185-64341

; Sequence 64341, Application US/10750185
 ; Publication No. US20050260603A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MMI GENOMICS, INC.
 ; ATTORNEY: DENISE, Sue K.
 ; ATTORNEY: KERR, Richard
 ; ATTORNEY: ROSENFIELD, David
 ; ATTORNEY: HOLM, Tom
 ; ATTORNEY: BATES, Stephen
 ; ATTORNEY: FANTIN, Dennis
 ; FILE REFERENCE: MMII1100-2
 ; CURRENT APPLICATION NUMBER: US/10/750,185
 ; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
 ; FILE REFERENCE: MMII1100-2
 ; CURRENT APPLICATION NUMBER: US 60/437,482
 ; PRIOR APPLICATION NUMBER: US 60/437,482
 ; NUMBER OF SEQ ID NOS: 64922
 ; SOFTWARE: PatentIN version 3.1
 ; SEQ ID NO: 42464
 ; LENGTH: 2376

RESULT 5
 US-10-750-185-28759/c

; Sequence 28759, Application US/10750185
 ; Publication No. US20050260603A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MMI GENOMICS, INC.
 ; ATTORNEY: DENISE, Sue K.
 ; ATTORNEY: KERR, Richard
 ; ATTORNEY: ROSENFIELD, David
 ; ATTORNEY: HOLM, Tom
 ; ATTORNEY: BATES, Stephen
 ; ATTORNEY: FANTIN, Dennis
 ; FILE REFERENCE: MMII1100-2
 ; CURRENT APPLICATION NUMBER: US/10/750,185
 ; CURRENT FILING DATE: 2003-12-31
 ; PRIOR APPLICATION NUMBER: US 60/437,482
 ; PRIOR FILING DATE: 2002-12-31
 ; NUMBER OF SEQ ID NOS: 64922
 ; SOFTWARE: PatentIN version 3.1
 ; SEQ ID NO: 28759
 ; LENGTH: 1931
 ; TYPE: DNA
 ; ORGANISM: Bovine 19866881378846

Query Match Score 82.0%; DB 6; Length 1165;
 Best Local Similarity 94.4%; Pred. No. 71;
 Matches 17; Conservative 0; Mismatches 1;
 Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATGAGAAA 18
 Db 689 GAGTTGGAGGATGAGAAA 706

RESULT 6
 US-10-750-185-42464/c

; Sequence 42464, Application US/10750185
 ; Publication No. US20050260603A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MMI GENOMICS, INC.
 ; ATTORNEY: DENISE, Sue K.
 ; ATTORNEY: KERR, Richard
 ; ATTORNEY: ROSENFIELD, David
 ; ATTORNEY: HOLM, Tom
 ; ATTORNEY: BATES, Stephen
 ; ATTORNEY: FANTIN, Dennis
 ; FILE REFERENCE: MMII1100-2
 ; CURRENT APPLICATION NUMBER: US/10/750,185
 ; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
 ; FILE REFERENCE: MMII1100-2
 ; CURRENT APPLICATION NUMBER: US 60/437,482
 ; PRIOR APPLICATION NUMBER: US 60/437,482
 ; NUMBER OF SEQ ID NOS: 64922
 ; SOFTWARE: PatentIN version 3.1
 ; SEQ ID NO: 42464
 ; LENGTH: 2376


```

Db 1352 TGGAAAGGTAGAACG 1336
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; GENERAL INFORMATION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: PastSEQ for Windows Version 4.0
; SEQ ID NO: 14464
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-995-561-14464

RESULT 11
US-10-750-185-40303/c
; Sequence 40303, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFIELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMII1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO: 40303
; LENGTH: 3065
; TYPE: DNA
; ORGANISM: Bovine 19866880368159
; US-10-750-185-40303

Query Match 77.0%; Score 15.4%; DB 6; Length 3065;
Best Local Similarity 94.1%; Pred. No. 2.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 AGTAGGAAGGTAGAACG 18
Db 455 AGTAGGAAGGTAGAACG 439

RESULT 12
US-10-995-561-13293
; Sequence 13293, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: PastSEQ for Windows Version 4.0
; SEQ ID NO: 13293
; LENGTH: 645179
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-995-561-13293

Query Match 77.0%; Score 15.4%; DB 6; Length 645179;
Best Local Similarity 94.1%; Pred. No. 2.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GAGTAGGAGGTAGAACG 17
Db 547755 GAGAGGAGGTAGAACG 547771

RESULT 13
US-10-995-561-14464/c
; Sequence 14464, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMII1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; US-10-750-185-48039/c
; Sequence 48039, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFIELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMII1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIORITY APPLICATION NUMBER: US 60/437,482
; PRIORITY FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO: 48039
; LENGTH: 687
; TYPE: DNA
; ORGANISM: Bovine 19866881734474
; US-10-750-185-48039

Query Match 76.0%; Score 15.2%; DB 6; Length 687;
Best Local Similarity 85.0%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GAGTAGGAGGTAGAACG 20
Db 74 GAGTGGAGGGTAGAGCG 55

RESULT 15
US-10-750-185-29945
; Sequence 29945, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFIELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMII1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; US-10-750-185-29945
; Sequence 29945, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFIELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMII1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185

```

; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO: 29945
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Bovine 19866880773646
; US-10-750-185-29945

Query Match 76.0%; Score 15.2; DB 6; Length 1400;
Best Local Similarity 85.0%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAACG 20
Db 980 GAGAAGGAAGAAAGAAAG 999

Search completed: January 2, 2006, 15:23:18
Job time : 165 secs

THIS PAGE IS BLANK

GenCore version 5.1.6
 Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 2, 2006, 13:57:00 ; Search time 2130 Seconds
 Perfect score: 20 (without alignments)
 Sequence: 533.741 Million cell updates/sec

Title: US-09-675-650-4
 Scoring table: IDENTITY NUC
 Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database : GenEmbl:*

```
1: 9b_ba:*
2: 9b_in:*
3: 9b_env:*
4: 9b_cm:*
5: 9b_cv:*
6: 9b_pat:*
7: 9b_ph:*
8: 9b_pr:*
9: 9b_ro:*
10: 9b_sts:*
11: 9b_sy:*
12: 9b_un:*
13: 9b_vl:*
14: 9b_hrg:*
15: 9b_pl:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	6 AX107860	AX107860 Sequence
c 2	20	100.0	506	6 COB55951	COB55951 Sequence
c 3	20	100.0	506	6 AX107857	AX107857 Sequence
c 4	20	100.0	535	8 AF103908	AF103908 Homo sapi
c 5	20	100.0	128480	8 AL359314	AL359314 Human DNA
c 6	19	95.0	237224	14 AC141860	AC141860 Gallus ga
c 7	18	90.0	130365	14 AC150126	AC150126 Gallus ga
c 8	18	90.0	200558	14 AC150067	AC150067 Gallus ga
c 9	18	90.0	262093	14 AC150039	AC150039 Gallus ga
c 10	17.4	87.0	58609	14 AC142494	AC142494 Ratetus no
c 11	17.4	87.0	130244	14 AC013518	AC013518 Homo sapi
c 12	17.4	87.0	130336	8 HS17K7	AL035078 Human DNA
c 13	17.4	87.0	153206	9 AC102304	AC102304 Mus muscu
c 14	17.4	87.0	153305	14 AC139606	AC139606 Ratetus no
c 15	17.4	87.0	162197	9 AC139333	AC139333 Mus muscu
c 16	17.4	87.0	163280	14 AC128280	AC128280 Ratetus no
c 17	17.4	87.0	166525	8 AC022413	AC022413 Homo sapi
c 18	17.4	87.0	170028	5 CR792439	CR792439 Zebrafish

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RESULT 1
 AX107860 LOCUS Sequence 4 From Patent WO0123550.
 DEFINITION Sequence 4 From Patent WO0123550.
 ACCESSION AX107860
 VERSION GI:13923251
 KEYWORDS synthetic construct
 SOURCE synthetic construct
 ORGANISM synthetic construct
 other sequences; artificial sequences.

REFERENCE 1 Busse, U., Chypre, C. and Fradet, Y.
 TITLE Pca3 messenger rna species in benign and malignant prostate tissues
 AUTHORS Busse, U., Chypre, C. and Fradet, Y.
 JOURNAL Diagnocure Inc. (CA)
 PATENT WO 0123550-A 4 05-APR-2001;
 FEATURES Source
 1. 20
 /organism="Synthetic construct"
 /mol type="unassigned DNA"
 /db_xref="taxon:31630"
 /note="Oligonucleotide"

ORIGIN
 Query Match 100.0%; Score 20; DB 6; Length 20;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 LOCUS CQ855951 Sequence 7 From Patent WO2004070056.
 DEFINITION CQ855951
 ACCESSION CQ855951
 VERSION CQ855951.1 GI:51850785
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

Eu^{karyota}; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

REFERENCE 1
AUTHORS Fradet, Y., Chypre, C., Piché, L. and Garon, G.
TITLE Method to detect prostate cancer in a sample
JOURNAL Patent : WO 200407056-A 7 19-AUG-2004;
FEATURES Diagnocure Inc. (CA)
Source Location/Qualifiers 1. .506
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:3606"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 506;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTGGAAAGGATGAAACG 20
Db 111 GAGTAGGAAAGGATGAAACG 92

RESULT 3

AX107857/C AX107857 Sequence 1 from Patent WO0123550. 506 bp DNA linear PAT 30-APR-2001

LOCUS DEFINITION Sequence 1 from Patent WO0123550.
ACCESSION AX107857
VERSION AX107857.1 GI:13923249

KEYWORDS Homo sapiens (human)

SOURCE ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

REFERENCE 1
AUTHORS Busse, U., Chypre, C. and Fradet, Y.
TITLE PCA3 messenger rna species in benign and malignant prostate tissues
JOURNAL Patent : WO 0123550-A 1 05-APR-2001;
Diagnocure Inc. (CA)

FEATURES Location/Qualifiers 1. .506
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 506;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAAGGATGAAACG 20
Db 111 GAGTAGGAAAGGATGAAACG 92

RESULT 4

AF103908/C AF103908 Homo sapiens non-coding RNA DD3 gene, exons 2, 3, and 4. 5435 bp DNA linear PRI 14-AUG-2000

ORGANISM Eu^{karyota}; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

KEYWORDS Homo sapiens (human)

REFERENCE 1
AUTHORS Busselmakers, M.J., van Bokhoven, A., Verhaegh, G.W., Smit, F.P., Karthaus, H.F.M., Schalken, J.A., Debruyne, F.M., Ru,N. and Isaacs,W.B.

TITLE DD3: a new prostate-specific gene, highly overexpressed in prostate

cancer Res. 59 (23), 5975-5979 (1999)

JOURNAL Cancer Res.

PUBMED 1086244

REFERENCE 2 (bases 1 to 5435)

AUTHORS Busselmakers, M.J.G., Van Bokhoven, A., Verhaegh, G.W., Smit, F.P., Karthaus, H.F.M., Schalken, J.A., Debruyne, F.M.J., Ru,N. and Isaacs,W.B.

TITLE Direct Submission

JOURNAL Submitted (28-OCT-1998) Urology Research Laboratory, University Hospital Nijmegen, P.O. Box 9101, Nijmegen 6500 HB, Netherlands

FEATURES Location/Qualifiers 1. .5435
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/map="9q11-q22"
repeat_region 1..78
/rpt_family="AluY"
exon 533..697
/note="alternative exon present in 5 of cDNA clones"
/number=2
repeat_region 1035..1194
/rpt_family="AluY"
/rpt_type=dispersed
mRNA join(1571..1753,181..5435)
/product="non-coding RNA DD3"
/note="transcript III"
mRNA join(1571..1753,181..3579)
/product="non-coding RNA DD3"
/note="transcript (major) III"
mRNA join(1571..1753,1981..2517)
/product="non-coding RNA DD3"
/note="transcript I"
exon 1571..1753
/number=3
exon 1981..5435
/number=4
polyA_signal 2495..2499
/note="transcript I"
polyA_site 2517
/note="transcript I"
polyA_signal 3553..3558
/note="transcript (major) III"
polyA_site 3579
/note="transcript I"
repeat_region 5268..5423
/note="LINE"
/rpt_family="L1"
/rpt_type=dispersed
polyA_site 5435
/note="transcript III"

ORIGIN

Query Match 100.0%; Score 20; DB 8; Length 5435;

Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAAGGATGAAACG 20
Db 1838 GAGTAGGAAAGGATGAAACG 1819

RESULT 5

AL359314/C AL359314 Human DNA sequence from clone RP11-108L4 on chromosome 9 Contains the 5' end of the KIAA0367 gene and the 5' end of the PCA3 gene for prostate cancer antigen 3 (DD3). complete sequence.

DEFINITION The 5' end of the KIAA0367 gene and the 5' end of the PCA3 gene for prostate cancer antigen 3 (DD3).

ACCESSION AL359314
VERSION AL359314..14
KEYWORDS HTG; DD3; KIAA0367; PCA3; prostrate cancer.

ABDMIVYLNGATPERRMPGILGWMKCYQMDRRLRKNLKFIIYHPSWFTIRTLAVT
REPISSKPSSKKTKYNS"

CDS
join (complement (AL161626 .20 :182245 . .182224),
complement (AL161626 .20 :179205 . .179405),
complement (AL161626 .20 :171461 . .171634),
complement (AL161626 .20 :168702 . .168702),
complement (AL161626 .20 :164908 . .165010),
complement (AL161626 .20 :164200 . .164271))
/gene="XIAA0367"
/locus_tag="RP11-214N16.3-001"
/standard_name="OTTHUDNP0000021506"
/codon_start=1
/product="XIAA0367"
/protein_id="CAII0913_1"
/db_xref="GI:15859642"
/db_xref="InterPro:IPR001251"
/db_xref="UniProt/TREMBL:Q5T474"
/translation="MLKSCRSRASFSPSVRKPKLAPNTINSLDQSRTSILDDNLSPDEIDINYDEL
ADMREPEPNSSLNDTHPRRTRKLTAPNTINSLDQSRTSILDDNLSPDEIDINYDEL
MVKIEPVRVISGGDGGYGGDNATIVFACFLPPSSRADYHTMENLFLYVSTL
ELMVAEDEMYIPLNGATP"

gene
join (complement (47725 .>53531), complement (46301 . .47063),
complement (AL161626 .20 :182145 . .182224),
complement (AL161626 .20 :179205 . .179405),
complement (AL161626 .20 :171461 . .171631),
complement (AL161626 .20 :168702 . .165010),
complement (AL161626 .20 :164308 . .164271),
complement (AL161626 .20 :164140 . .164271),
complement (AL161626 .20 :153193 . .163279),
complement (AL161626 .20 :155914 . .156012),
complement (AL161626 .20 :141011 . .141322))
/gene="XIAA0367"
/locus_tag="RP11-214N16.3-006"
/join (complement (47725 .>53531), complement (46301 . .47063),
complement (AL161626 .20 :182145 . .182224),
complement (AL161626 .20 :179205 . .179405),
complement (AL161626 .20 :171461 . .171631),
complement (AL161626 .20 :168702 . .165010),
complement (AL161626 .20 :164308 . .164271),
complement (AL161626 .20 :164140 . .164271),
complement (AL161626 .20 :153193 . .163279),
complement (AL161626 .20 :155914 . .156012),
complement (AL161626 .20 :141011 . .141322))
/gene="XIAA0367"
/locus_tag="RP11-214N16.3-006"

mRNA
Query Match 100 0% Score 20 DB 8 Length 128480;
Matches 20; Similarity 100.0%; Pred. No. 14;
Definition 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAACG 20
Db 126936 GAGTAGGAAGGATAGAACG 126917

RESULT 6
AC141860 LOCUS AC141860 237234 bp DNA linear HTG 03-APR-2003
DEFINITION Gallus gallus clone TAM31-57D16, WORKING DRAFT SEQUENCE, 10 ordered
ACBSESSION pieces.
VERSION AC141860_2 GI:29501863
KEYWORD HTGS PHASE2; HTGS DRAFT
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus;
REFERENCE 1 (bases 1 to 237234)
AUTHORS Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,

Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J., Haghichi,P., Han,J., Hansen,N., Ho,L., Idol,J.R., Karins,E., Laric,P., Lee,Lin,S.-Q., Legasp,R., Maduro,Q.I., Maduro,V.B., Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.J., Pagnirigan,C., Pearson,R., Portnoy,M.E., Prasad,A., Reddi-Bougu,N., Schandler,K., Thomas,P.J., Touchman,J.W., Vogt,J.L., Staunstrup,S., Thomas,J.W., Thomas,M.G., Sison,C., Wetherby,K.D., Wiggin,L., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
TITLE
JOURNAL
REFERENCE 2 (bases 1 to 237234)
AUTHORS Green,E.D.
TITLE Direct Submission (19-MAR-2003) NIH Intramural Sequencing Center, 8717
JOURNAL
COMMENT 3 (bases 1 to 237234)
REFERENCE
AUTHORS Green,E.D.
TITLE Direct Submission (03-APR-2003) NIH Intramural Sequencing Center, 8717
JOURNAL
COMMENT Submitted (03-APR-2003) NIH Intramural Sequencing Center, 8717 USA
On Apr 3, 2003 this sequence version replaced gi:29124110.
----- Project Information
Center project name: dft
Center clone name: 057016
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@hgri.nih.gov
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 235353 bases at least Q40
Consensus quality: 235850 bases at least Q30
Insert size: 180000; 236231 bases at least Q20
Insert size: 236334; sum-of-contigs
Quality coverage: 8.83x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 12817: contig of 12817 bp in length
* 12818: 12917: gap of unknown length
* 12918: 85073: contig of 72156 bp in length
* 85173: gap of unknown length
* 85174: 88714: contig of 3541 bp in length
* 88715: 88814: gap of unknown length
* 88815: 113928: contig of 25114 bp in length
* 113929: 114028: gap of unknown length
* 114029: 191667: contig of 77639 bp in length
* 191668: gap of unknown length

* 191768 200289: contig of 8522 bp in length
 * 200290 200389: gap of unknown length
 * 200390 211976: contig of 11587 bp in length
 * 211976 212076: gap of unknown length
 * 212077 214145: contig of 2069 bp in length
 * 214146 214246: contig of 9949 bp in length
 * 214246 224194: gap of unknown length
 * 224195 224294: gap of unknown length
 * 224295 237334: Contig of 12940 bp in length
 FEATURES Location/Qualifiers
 source 1..237234
 /note="Gallus gallus"
 /mol type="Genomic DNA"
 /db Xref="Saxxon:9031"
 /Clone="TAM31-57D16"
 /clone lib="TAM31"
 misc_feature 1..12817
 /note="assembly_fragment
 clone end:T7
 vector_side:left"
 12118..12117
 /estimated_length=unknown
 misc_feature 12918..85073
 /note="assembly_fragment"
 85074..85173
 /estimated_length=unknown
 misc_feature 85174..88714
 /note="assembly_fragment"
 88715..88914
 /estimated_length=unknown
 88815..111928
 /note="assembly_fragment"
 113929..114028
 /estimated_length=unknown
 114029..1151667
 /note="assembly_fragment"
 191668..191767
 /estimated_length=unknown
 misc_feature 191768..201289
 /note="assembly_fragment"
 200290..200389
 /estimated_length=unknown
 200390..211976
 /note="assembly_fragment"
 211977..212076
 /estimated_length=unknown
 misc_feature 212077..214145
 /note="assembly_fragment"
 214146..214245
 /estimated_length=unknown
 misc_feature 214246..224194
 /note="assembly_fragment"
 224195..224294
 /estimated_length=unknown
 misc_feature 224295..237334
 /note="assembly_fragment
 clone end:SP6
 vector_side:right"
 ORIGIN
 Query Match 95.0%; Score 19; DB 14; Length 237234;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 ORIGIN
 RESULT 7
 AC150126 AC150126 130365 bp DNA linear HTG 01-JUL-2004
 LOCUS Gallus gallus clone WAG-32P6 WORKING DRAFT SEQUENCE 2 ordered
 DEFINITION Gallus gallus clone WAG-32P6 WORKING DRAFT SEQUENCE 3 ordered
 DEPICTION Gallus gallus clone CH261-73M16, WORKING DRAFT SEQUENCE 3 ordered
 pieces.
 AC15026 AC15026.1 GI:49533721
 VERSION HTG; HTGS PHASE2; HTGS DRAFT; HTGS_ACTIVEFIN.
 KEYWORDS SOURCE
 Gallus gallus (chicken)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus
 1 (bases 1 to 130365)
 REFERENCE
 AUTHORS DOB Joint Genome Institute.
 TITLE Unpublished
 JOURNAL
 2 (bases 1 to 130365)
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 DOB Joint Genome Institute.
 Direct Submission
 Submitted (01-JUL-2004) Production Genomics Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA
 94598-1698, USA
 COMMENT -----
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: http://www.jgi.doe.gov

 Project Information
 Center Project Name: 23633925
 Center Clone name: Chk_32P6

 Summary Statistics
 Consensus quality: 129752 bases at least Q40
 Consensus quality: 129883 bases at least Q30
 Consensus quality: 130130 bases at least Q20
 Estimated insert size: 138000; agarose-fp estimation
 Estimated insert size: 13265; sum-of-contigs estimation
 Quality coverage: 31.45 in Q20 bases; agarose-fp estimation
 Quality coverage: 33.32 in Q20 bases; sum-of-contigs estimation.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 56305: contig of 56305 bp in length
 * 56306 56405: gap of unknown length
 * 56406 130365: contig of 73960 bp in length.
 FEATURES
 Source
 /organism="Gallus gallus"
 /mol type="genomic DNA"
 /db Xref="taxon:9031"
 /clone lib="Texas A&M Wageningen Chicken BAC Library"
 gap
 2 AGTAGGAAGGATAGAAC 19
 /clone lib="Texas A&M Wageningen Chicken BAC Library"
 gap
 56306..56405
 /estimated_length=unknown
 ORIGIN
 Query Match 90.0%; Score 18; DB 14; Length 130365;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 ORIGIN
 RESULT 8
 AC150067 AC150067 200558 bp DNA linear HTG 01-JUL-2004
 LOCUS Gallus gallus clone CH261-73M16, WORKING DRAFT SEQUENCE 3 ordered
 DEFINITION pieces.
 DEPICTION

ACCESSION AC150067
 VERSION HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.
 SOURCE Gallus gallus (chicken)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauvia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
 REFERENCE 1. (bases 1 to 200558)
 AUTHORS DOE Joint Genome Institute.
 TITLE Unpublished
 JOURNAL 2. (bases 1 to 200558)
 REFERENCE 2. (bases 1 to 200558)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (01-JUL-2004) Production Genomics Facility, DOE Joint Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698 USA
 COMMENT ---G-Name Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: <http://www.jgi.doe.gov>
 Project Information
 Center Project Name: 2959111
 Center clone name: JF2-117B18
 Summary Statistics
 Consensus quality: 198774 bases at least Q40
 Consensus quality: 199779 bases at least Q30
 Consensus quality: 200130 bases at least Q20
 Estimated insert size: 200358; agarose-fp estimation
 Estimated insert size: 200358; sum-of-contigs estimation
 Quality coverage: 10.32 in Q20 bases; agarose-fp estimation
 Quality coverage: 10.3 in Q20 bases; sum-of-contigs estimation
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced.
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1. 13044: contig of 13044 bp in length
 * 13045 13144: gap of unknown length
 * 11745 11745: contig of 104439 bp in length
 * 117584 117583: gap of unknown length
 * 117684 200558: contig of 82875 bp in length.
 FEATURES source
 1. .200558
 /organism="Gallus gallus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9031"
 /clone="CH261-73M16"
 /clone.lib="CHORI-261 Chicken BAC Library"
 13045 ..13144
 /estimated_length=unknown
 117584 ..117583
 /estimated_length=unknown
 ORIGIN
 Query Match 90.0%; Score 18; DB 14; Length 200558;
 Best Local Similarity 100.0%; Pred. No. 1. 4e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Gaps 0
 gap 2 AGTAGGAAGGATAGAAC 19
 gap 88846 AGTAGGAAGGATAGAAC 88863
 ORIGIN
 Query Match 90.0%; Score 18; DB 14; Length 262093;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Gaps 0
 Db 173152 AGTAGGAAGGATAGAAC 173135
 RESULT 10
 AC142494
 LOCUS AC142494
 AC150039/c
 RESULT 9
 AC150039/c

DEFINITION	Rattus norvegicus clone CH230-164M14, *** SEQUENCING IN PROGRESS
ACCESSION	AC142494
VERSION	AC142494.1
KEYWORDS	GI :29423811
SOURCE	HTG; HTGS; PHASEI.
ORGANISM	Rattus norvegicus
REFERENCE	Muzny, D., Marie,, Metzger, M., Lee,, Abramzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angniano, D., Anvallebechhi V., Aoyagi, A., Ayodebi, M., Baca, E., Barber, M., Barnstead, M., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, J., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chu, J., Cleveland, C., Cockrell, R., Cox, J., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, I., De Anda, C., Dederich, D., Delgado, O., Denson, S., Derano, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durval, B., Eaves, K., Egan, A., Escott, M., Eugene, C., Evans, C.A., Failes, T., Fan, G., Fernandes, S., Finley, M., Flagg, N., Forbes, J., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebara, E., Geer, K., Gill, R., Grady, M., Guerrra, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hodgeson, A., Hogenes, M., Hollins, B., Howell, S., Huijk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, J., Kovar, C., Kowalewski, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Lounges, H., Lorado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindarne, M., Mahmoud, M., Malloy, K., Margum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M., McNeil, T., Meenend, E., Milosavljevic, A., Miner, G., Mitja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakelume, O., Okononu, G., Olarinpunsgaon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sission, I., Sitter, C.D., Smais, D., Snead, A., Sodergren, E., Song, X., Sorelle, R., Sose, J., Steinle, M., Strong, R., Surtron, A., Svatek, A., Tabor, P., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, J., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 58609)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (01-APR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	Genome Center Center: Baylor College of Medicine Center code: BCM Web site: http://www.hgsc.bcm.tmc.edu/ Contact: hgsc-help@bcm.tmc.edu
TITLE	Project Information

Center project name: KBPD
 Center clone name: CH230-164M14
 ----- Summary
 AC142494
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap; version 0_990329
 Consensus quality: 45901 bases at least Q30
 Consensus quality: 50564 bases at least Q30
 Consensus quality: 53610 bases at least Q20
 Estimated insert size: 48044; sum-of-contigs estimation
 Quality coverage: 1x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank/draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 34 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1284: contig of 1284 bp in length
 * 1 1285: 1384: gap of unknown length
 * 1 1385: 2547: contig of 1163 bp in length
 * 1 2548: 2647: gap of unknown length
 * 1 2648: 3743: contig of 1096 bp in length
 * 1 3744: 3843: gap of unknown length
 * 1 3844: 5041: contig of 1198 bp in length
 * 1 5042: 5141: gap of unknown length
 * 1 5142: 6213: contig of 1072 bp in length
 * 1 6214: 6313: gap of unknown length
 * 1 6314: 7364: contig of 1051 bp in length
 * 1 7365: 7464: gap of unknown length
 * 1 7465: 8764: contig of 1300 bp in length
 * 1 8765: 8864: gap of unknown length
 * 1 8865: 10214: contig of 1450 bp in length
 * 1 10214: 10414: gap of unknown length
 * 1 10414: 111609: contig of 1195 bp in length
 * 1 111610: 11709: gap of unknown length
 * 1 11710: 13017: contig of 1308 bp in length
 * 1 13018: 13117: gap of unknown length
 * 1 13118: 14275: contig of 1258 bp in length
 * 1 14275: 14376: gap of unknown length
 * 1 14376: 15744: contig of 1269 bp in length
 * 1 15745: 15844: gap of unknown length
 * 1 15845: 17087: contig of 1243 bp in length
 * 1 17088: 17187: gap of unknown length
 * 1 17188: 18458: contig of 1271 bp in length
 * 1 18458: 18558: gap of unknown length
 * 1 18558: 20095: contig of 1537 bp in length
 * 1 20095: 20299: gap of unknown length
 * 1 20299: 21794: contig of 1599 bp in length
 * 1 21795: 21894: gap of unknown length
 * 1 21895: 23155: contig of 1261 bp in length
 * 1 23155: 24223: gap of unknown length
 * 1 24223: 24524: gap of unknown length
 * 1 24524: 25337: contig of 1314 bp in length
 * 1 25338: 26038: 28927: contig of 1990 bp in length
 * 1 26038: 28928: 28927: gap of unknown length
 * 1 28928: 29999: gap of unknown length
 * 1 29999: 31988: gap of unknown length
 * 1 31988: 31499: gap of unknown length
 * 1 31499: 34550: contig of 2652 bp in length
 * 1 34550: 34251: gap of unknown length
 * 1 34251: 36022: contig of 1672 bp in length
 * 1 36022: 36123: gap of unknown length
 * 1 36123: 37639: contig of 1567 bp in length
 * 1 37639: 37789: gap of unknown length
 * 1 37789: 37790: contig of 2211 bp in length

11150 11249: gap of 100 bp
* * 12019: contig of 770 bp in length
11250 12119: gap of 100 bp
* * 12120: contig of 741 bp in length
12020 12860: gap of 100 bp
* * 12861: contig of 752 bp in length
12961 13712: gap of 100 bp
* * 13713: contig of 762 bp in length
13813 14574: gap of 100 bp
* * 14575: contig of 769 bp in length
14675 15443: gap of 100 bp
* * 15544: contig of 746 bp in length
15544 16289: gap of 100 bp
* * 16290: contig of 781 bp in length
16390 17170: gap of 100 bp
* * 17171: contig of 762 bp in length
17270 18032: gap of 100 bp
* * 18132: contig of 771 bp in length
18133 18903: gap of 100 bp
* * 18904: contig of 792 bp in length
18904 19795: gap of 100 bp
* * 19796: contig of 759 bp in length
19896 20634: gap of 100 bp
* * 20754: contig of 744 bp in length
20755 21498: gap of 100 bp
* * 21499: contig of 768 bp in length
21599 22366: gap of 100 bp
* * 22367: contig of 766 bp in length
22466 23232: gap of 100 bp
* * 23332: contig of 769 bp in length
23333 24101: gap of 100 bp
* * 24102: contig of 710 bp in length
24002 24911: gap of 100 bp
* * 24912: contig of 782 bp in length
25012 25733: gap of 100 bp
* * 25894: contig of 781 bp in length
25894 26674: gap of 100 bp
* * 26675: contig of 750 bp in length
26754: contig of 710 bp in length
2675 27624: gap of 100 bp
* * 27625: contig of 777 bp in length
28402 28501: gap of 100 bp
* * 28502: contig of 756 bp in length
29358 29357: gap of 100 bp
* * 29358: contig of 779 bp in length
30137 30236: gap of 100 bp
* * 30237: contig of 762 bp in length
30999 31098: gap of 100 bp
* * 31098: contig of 760 bp in length
31099 31188: gap of 100 bp
* * 31188: contig of 781 bp in length
31959 31958: gap of 100 bp
* * 31959: contig of 783 bp in length
32742 32841: gap of 100 bp
* * 32842: contig of 758 bp in length
33600 33698: gap of 100 bp
* * 33698: contig of 781 bp in length
33700 34480: gap of 100 bp
* * 34481: contig of 764 bp in length
34534: contig of 100 bp
* * 35245: gap of 100 bp
* * 35245: contig of 758 bp in length
35445: gap of 100 bp
* * 36202: contig of 762 bp in length
36303: gap of 100 bp
* * 37026: contig of 724 bp in length
37027 37124: gap of 100 bp
* * 37127: contig of 784 bp in length
37911 38010: gap of 100 bp
* * 38011: contig of 765 bp in length
38873 38872: gap of 100 bp
* * 38873: contig of 775 bp in length
39648 39747: gap of 100 bp
* * 39748: contig of 767 bp in length
40514: gap of 100 bp
* * 40515: contig of 765 bp in length
41380 41479: gap of 100 bp
* * 41480: contig of 757 bp in length
42237 42336: gap of 100 bp

* 42337 43080: contig of 744 bp in length
* * 43081 43180: gap of 100 bp
* * 43181: contig of 782 bp in length
* * 43962: gap of 100 bp
* * 44062: contig of 775 bp in length
* * 44063 44837: gap of 100 bp
* * 44937: contig of 786 bp in length
* * 44938 45723: contig of 780 bp
* * 45724 45824: gap of 100 bp
* * 45824: contig of 784 bp in length
* * 46608 46707: gap of 100 bp
* * 46708: contig of 786 bp in length
* * 47494 47593: gap of 100 bp
* * 47594 48327: contig of 734 bp in length
* * 48328 48427: gap of 100 bp
* * 48428 49171: contig of 744 bp in length
* * 49172 49271: gap of 100 bp
* * 49272: contig of 747 bp in length
* * 50018: gap of 100 bp
* * 50019 50118: gap of 100 bp
* * 50119 50886: contig of 768 bp in length
* * 50887 50986: gap of 100 bp
* * 50987 51755: contig of 769 bp in length
* * 51756 51855: gap of 100 bp
* * 51856: contig of 778 bp in length
* * 52633: gap of 100 bp
* * 52634 52734: contig of 775 bp in length
* * 53508 53509: gap of 100 bp
* * 53509 53608: gap of 100 bp
* * 53609 54371: contig of 763 bp in length
* * 54372 54471: gap of 100 bp
* * 54472 55232: contig of 761 bp in length
* * 55233: gap of 100 bp
* * 55234 55332: contig of 775 bp in length
* * 55333 56121: contig of 789 bp in length
* * 56122: gap of 100 bp
* * 56222 56222: contig of 751 bp in length
* * 56973 57072: gap of 100 bp
* * 57073 57848: contig of 776 bp in length
* * 57849 57948: gap of 100 bp
* * 57949 58698: contig of 750 bp in length
* * 58699 58798: gap of 100 bp
* * 58799 59668: gap of 100 bp
* * 59669 60436: contig of 768 bp in length
* * 60437 60536: gap of 100 bp
* * 60537 61302: contig of 766 bp in length
* * 61303 61402: gap of 100 bp
* * 61403 62167: contig of 765 bp in length
* * 62168 62267: gap of 100 bp

Query Match 87.0%; Score 17.4%; DB 14; Length 130244;
Best Local Similarity 94.7%; Pred. No. 3e+02; Indels 0; Gaps 0;

LOCUS HS17K7 Human DNA sequence from clone RP1-17K7 on chromosome 11p13,
DEFINITION complete sequence.
ACCESSION AL035078
KEYWORDS HPG.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo
VERSION 34364
REFERENCE 1 (bases 1 to 130336)
AUTHORS Martin S.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,

COMMENT	Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Aug 30, 1999 this sequence version replaced gi_5804069. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Embl: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP1-17K7 is from the library RPCI-1 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm	AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,T., LaRocque,K., Lamasares,R., Landers,T., Lenoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marcus,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheevers,R., Meldrim,J., Meneus,L., Minova,T., Mienga,V., Murphy,T., Naytor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Riley,R., Rose,C., Rogov,P., Schupback,R., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Seaman,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.J., Tesfaye,S., Theodore,J., Topham,K., Vo,T., Travers,M., Travis,R., Trujillo,J., Vassiliev,H., Vieil,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A., and Zody,M.	TITLE JOURNAL Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 300 Charles Street, Cambridge, MA 02141, USA 3 (bases 1 to 153206) REFERENCE Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Halez,N., Hagogian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbutt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Menesu,L., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rize,C., Rogov,P., Roman,J., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J.J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Vieil,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A., and Zody,M.
FEATURES	source 1 .. 130336 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="RGD:RGICIP704K0717" /db_xref="taxon:9606" /chromosome="11" /map="p13" /clone="RP1-17K7" /clone_lib="RPCI-1" misc_feature 1 /note="Clone_left_end: RP1-17K7" misc_feature 20399 /note="Clone_right_end: RP1-65P5" /note="Clone_right_end: RP1-17K7" /note="N1329 bases of Trn10 (J01859) removed here. This sequence represents the duplicated flanking sequence of the Trn10." 130336 /note="Clone_right_end: RP1-17K7" misc_feature ORIGIN	AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Halez,N., Hagogian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbutt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Menesu,L., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rize,C., Rogov,P., Roman,J., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J.J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Vieil,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A., and Zody,M.	TITLE JOURNAL Submitted (08-FEB-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 153206)
DEFINITION	1 GAGTAGGAAGGATAGAAC 19 Mus musculus chromosome 7, clone RP24-403P14, complete sequence.	AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Halez,N., Hagogian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbutt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Menesu,L., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rize,C., Rogov,P., Roman,J., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J.J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Vieil,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A., and Zody,M.	TITLE JOURNAL Submitted (08-FEB-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 153206)
AC102304	AC102304 LOCUS 153206 bp DNA linear ROD 10-MAR-2005 DEFINITION Mus musculus chromosome 7, clone RP24-403P14, complete sequence.	AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Halez,N., Hagogian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbutt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Menesu,L., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rize,C., Rogov,P., Roman,J., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J.J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Vieil,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A., and Zody,M.	TITLE JOURNAL Submitted (08-FEB-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 153206)
AC102304	AC102304 LOCUS 153206 bp DNA linear ROD 10-MAR-2005 DEFINITION Mus musculus chromosome 7, clone RP24-403P14, complete sequence.	AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Halez,N., Hagogian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbutt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Menesu,L., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rize,C., Rogov,P., Roman,J., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J.J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Vieil,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A., and Zody,M.	TITLE JOURNAL Submitted (08-FEB-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 153206)
KEYWORDS	HTG.	AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Halez,N., Hagogian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbutt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Menesu,L., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rize,C., Rogov,P., Roman,J., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J.J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Vieil,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A., and Zody,M.	TITLE JOURNAL Submitted (08-FEB-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 153206)
SOURCE	Mus musculus (house mouse)	AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Halez,N., Hagogian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbutt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Menesu,L., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rize,C., Rogov,P., Roman,J., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J.J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Vieil,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A., and Zody,M.	TITLE JOURNAL Submitted (08-FEB-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 153206)
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Muridae; Murinae; Mus.	AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Halez,N., Hagogian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbutt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Menesu,L., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rize,C., Rogov,P., Roman,J., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J.J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Vieil,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A., and Zody,M.	TITLE JOURNAL Submitted (08-FEB-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 153206)
REFERENCE	1 (bases 1 to 153206)	AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Halez,N., Hagogian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbutt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Menesu,L., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rize,C., Rogov,P., Roman,J., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J.J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Vieil,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A., and Zody,M.	TITLE JOURNAL Submitted (08-FEB-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 153206)

JOURNAL Submitted (10-MAR-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT On Mar 10, 2005 this sequence version replaced gi:58743554.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://Ftp.Genome.Washington.Edu/RM/RepeatMasker.html>
 Genome Center
 Center: Broad Institute of MIT and Harvard
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@broad.mit.edu
 Project Information
 Center project name: Li8328
 Center clone name: 403_P_14

FEATURES source Some of the sequence contained within base pairs 118748 to the end of the clone was stolen from accession AC127696.
 Location.Qualifiers

repeat_region 1.
 /clone="153206" /organism="Mus musculus"
 /mol_type="Genomic DNA!" /db_xref="taxon:10090"
 /chromosome="7" /map="7"
 /clone="RP2-403P14" /clone_id="RPCI-24_Male_Mouse_BAC"
 repeat_region 940..1111 /rpt_family="Lx6"
 repeat_region 1117..1177 /rpt_family="GA-rich"
 repeat_region 1180..1277 /rpt_family="Lx7"
 repeat_region 1286..1439 /rpt_family="L1M2"
 repeat_region 1562..13701 complement(1548..1738)
 repeat_region 1804..2028 /rpt_family="L1MC3"
 repeat_region /rpt_family="B3A" complement(2735..27383)
 repeat_region /rpt_family="MT2B" Query Match 87.0%; Score 17.4; DB 9; Length 153206;
 repeat_region 3562..3701 /rpt_family="B1_MM" Best Local Similarity 94.7%; Pred. No. 3e+02;
 repeat_region /rpt_family="MT2B" Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 repeat_region 5745..5900 complement(5489..5744)
 repeat_region /rpt_family="L1" /rpt_family="B3C" AC139606 AC139606 Rattus norvegicus (Norway rat)
 repeat_region complement(5901..7161) AC139606 AC139606 Rattus norvegicus clone CH230-388D19, WORKING DRAFT SEQUENCE, 28
 repeat_region /rpt_family="BGLII" DEFINITION unorderd pieces.
 repeat_region complement(7577..7740) AC139606 AC139606 HTGS PHASE1; HTGS DRAFT
 repeat_region 7739..7886 KEYWORDS
 repeat_region /rpt_family="L1M4" SOURCE Rattus norvegicus
 repeat_region 7903..8920 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteostomi;
 repeat_region complement(3702..3870) Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 repeat_region /rpt_family="MT2B" REFERENCE Scuturognathus; Muridae; Murinae; Rattus.
 repeat_region 9787..10006 1 (bases 1 to 13305)
 repeat_region /rpt_family="Lx2" AUTHORS Muzny, D. Marie.; Netzker, M. Lee.; Abramzon, S.; Adams, C.; Alder, J.;
 repeat_region 10005..10564 Allen, H.; Alstroem, S.; Amm, A.; Angiano, D.;
 repeat_region /rpt_family="Lx2" Ayala-Bichi, V.; Aoyagi, A.; Ayodeji, M.; Baca, E.; Baden, H.;
 repeat_region 10365..10775 Baldwin, D.; Bandaranaike, D.; Barber, M.; Barnstead, M.; Benahmed, F.,
 repeat_region 9146..9748 Biswal, K.; Blair, J.; Blankenburg, K.; Blyth, P.; Brown, M.,
 repeat_region 10776..10998 Bryant, N.; Buhay, C.; Burch, P.; Burrell, K.; Calderon, E.,
 repeat_region /rpt_family="Lx2" Cardenas, V.; Carter, K.; Cavazos, J.; Caesar, H.; Center, A.;
 repeat_region 10951..10975 Chacko, J.; Chavez-D.; Chen, G.; Chen, R.; Chen, Y.; Chen, Z.; Chu, J.,
 repeat_region /rpt_family="AT_rich" Cleveland, C.; Cochrell, R.; Cox, C.; Coyle, M.; Cree, A.; D'Souza, L.,
 repeat_region 11056..11269 Davila, M.L.; Davis, C.; Davy-Carroll, L.; De Andrade, C.; Dederich, D.,
 repeat_region /rpt_family="Lx7"

Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Ezen,A., Escott,M., Eugene,C., Evans,C.A., Fall,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Georges,E., Geer,K., Gill,R., Gladys,M., Guerra,W., Guerra,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlik,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodges,M., Hollins,B., Howells,S., Hulky,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Kapathy,S., Kelly,S., Khan,Z., King,L., Kovar,C., Lewis,L., Lio,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenshew,L., Loulaed,H., Lozado,R.J., Lu,X., Ma,J., Naheshwari,M., Mahindarts,M., Mahmoud,M., Malloy,K., Mangum,A., Nangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Nawner,S., Mcleod,M., Mcneile,T., Meeden,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidas,M., Murphy,M., Nair,I., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwakelueh,O., Okuongu,G., Olarpunsagoon,A., Pal,S., Parks,K., Passernak,S., Paul,H., Perez,A., Perez,L., Pernoch,C., Plopper,F., Poidexter,A., Popovic,D., Primus,E., Pu,L.-J., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Soderren,B., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Thomas,N., Thomas,S., Tringe,A., Trejo,Z., Usmani,V., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,P., Williams,G., Wilson,R., Wleczky,R., Wooden,H., Worley,K., Wright,D., Wright,J., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zha,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE

Unpublished

2 (bases 1 to 153305)

Worley,K.C.

Direct Submission

Submitted (07-FEB-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

Worley,K.C.

Direct Submission

Submitted (14-APR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

Worley,K.C.

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-heip@bcm.tmc.edu

----- Project Information

Center project name: KDP

Center clone name: CH230-38BD19

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly program: Phrap; version 0.90329

Consensus quality: 140842 bases at least Q40

Consensus quality: 143014 bases at least Q30

Consensus quality: 145056 bases at least Q20

Estimated insert size: 14120; sum-of-contigs estimation

Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently consists of 28 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 1194: contig of 1194 bp in length
 * 1195 1294: gap of unknown length
 * 1295 2358: contig of 1064 bp in length
 * 2359 2458: gap of unknown length
 * 2459 3729: contig of 1271 bp in length
 * 3730 3829: gap of unknown length
 * 3830 4900: contig of 1071 bp in length
 * 4901 5000: gap of unknown length
 * 5001 6880: contig of 1880 bp in length
 * 6881 6980: gap of unknown length
 * 6981 8123: contig of 1143 bp in length
 * 8124 8223: gap of unknown length
 * 8224 9820: contig of 1597 bp in length
 * 9821 9920: gap of unknown length
 * 9921 11804: contig of 1884 bp in length
 * 11805 11904: gap of unknown length
 * 11905 12944: contig of 1040 bp in length
 * 12945 13044: gap of unknown length
 * 13045 14397: contig of 1353 bp in length
 * 14398 14497: gap of unknown length
 * 14498 16343: contig of 1846 bp in length
 * 16344 16443: gap of unknown length
 * 16444 17857: contig of 1414 bp in length
 * 17858 17957: gap of unknown length
 * 17958 19665: contig of 1708 bp in length
 * 19666 19765: gap of unknown length
 * 19765 23422: contig of 3657 bp in length
 * 23423 23522: gap of unknown length
 * 23523 26283: contig of 2761 bp in length
 * 26284 26383: gap of unknown length
 * 26384 30690: contig of 4307 bp in length
 * 30691 30790: gap of unknown length
 * 30790 35725: contig of 4935 bp in length
 * 35726 35825: gap of unknown length
 * 35826 40780: contig of 4955 bp in length
 * 40781 40880: gap of unknown length
 * 40881 46579: contig of 5693 bp in length
 * 46580 46679: gap of unknown length
 * 46680 54538: contig of 7859 bp in length
 * 54539 54638: gap of unknown length
 * 54639 66093: contig of 11455 bp in length
 * 66094 66193: gap of unknown length
 * 66194 76359: contig of 10166 bp in length
 * 76359 76459: gap of unknown length
 * 76460 86351: contig of 8982 bp in length
 * 86352 86451: gap of unknown length
 * 86452 96947: contig of 10496 bp in length
 * 96948 97047: gap of unknown length
 * 97048 107903: contig of 10856 bp in length
 * 107904 108003: gap of unknown length
 * 108004 116701: contig of 8698 bp in length
 * 116702 116801: gap of unknown length
 * 116802 131970: contig of 15169 bp in length
 * 131971 132070: gap of unknown length
 * 132071 153305: contig of 21235 bp in length.

FEATURES

source

1. 153305

/organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="Exxon10116"
 /clones="CH230-38BD19"
 1195..1294
 /estimated_length=unknown
 2359..2458
 /estimated_length=unknown

gap

gap

gap 3730 .3829 /estimated_length=unknown
gap 4901 5000 /estimated_length=unknown
gap 6881 .6980 /estimated_length=unknown
gap 8124 .8223 /estimated_length=unknown
gap 9821 .9220 /estimated_length=unknown
gap 11805 .11904 /estimated_length=unknown
gap 12945 .13044 /estimated_length=unknown
gap 14398 .14497 /estimated_length=unknown
gap 16344 .16443 /estimated_length=unknown
gap 17858 .17957 /estimated_length=unknown
gap 19666 .19765 /estimated_length=unknown
gap 23422 .23522 /estimated_length=unknown
gap 26284 .26383 /estimated_length=unknown
gap 30691 .30790 /estimated_length=unknown
gap 35726 .35825 /estimated_length=unknown
gap 40781 .40880 /estimated_length=unknown

Query Match Score 17.4%; DB 14; Length 133305;
Best Local Similarity 94.7%; Pred. No. 3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGTAGGAGGATAGAAC 19
Db 24999 GAGCAGGAGGATAGAAC 25017

RESULT 15
AC139333 LOCUS AC139333 162197 bp DNA linear ROD 28-JAN-2005
DEFINITION Mus musculus BAC clone RP24-83F12 from 3, complete sequence.
VERSION AC139333_4 GI:52077879
KEYWORDS HTG,
ORGANISM Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathia; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 162197)
AUTHORS Kanchi,K., Haglund,K., Bielicki,L. and Meyer,R.
TITLE The sequence of Mus musculus BAC clone RP24-83F12
JOURNAL Unpublished (2001)
PUBLISHER Wilson,R.K.
REFERENCE 2 (bases 1 to 162197)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (30-JAN-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 3 (bases 1 to 162197)
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (20-AUG-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 162197)
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

REFERENCE 5 (bases 1 to 162197)
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (28-JAN-2005) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT On Sep 15, 2004 this sequence version replaced gi:51468472.
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
Summary Statistics
Center project name: M_BB0083F12

NOTICE This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION: Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION: The RP1-24 BAC library has been constructed by Pieter de Jong and
coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen
and/or brain genomic DNA. The clone and detailed information can be
obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION: This sequence is the entire insert of the clone.

FEATURES Location/Qualifiers

1. 162197 /organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="3"
/map="3"
/clone.lib="RPCI-24"
/clone_id="RPCI-24_1870..18791"
/note="unresolved simple sequence repeat."
unsure
misc_feature 31786 ..31833 /note="Sequence derived from PCR product of project DNA."
unsure
32932 ..33006 /note="Sequence derived from one plasmid subclone."
unsure
33345 ..33460 /note="Sequence derived from one plasmid subclone."
unsure
34531 ..34576 /note="Sequence derived from one plasmid subclone."
unsure
125138 ..125349 /note="Sequence derived from one plasmid subclone."
unsure
125444 ..125457 /note="Sequence derived from one plasmid subclone."
ORIGIN

Query Match Score 87.0%; DB 9; Length 162197;
Best Local Similarity 94.7%; Pred. No. 3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGTGGAGGATAGAAC 19
Db 102773 GAGTGGAGGATAGAAC 102791

Tue Jan 3 12:12:12 2006

us-09-675-650-4.rge

Page 14

Search completed: January 2, 2006, 14:41:11
Job time : 2135 secs

November 2005

Published Applications Nucleic Acid and Published Applications Amino Acid database searches now generate two sets of results each. The Published Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published Applications New databases; older published applications make up the Published Applications Main databases.

Searches run against Nucleic Acid Published Applications produce two sets of results, with the extensions .rnpbm (Published Applications_NA_Main) and .rnpbm (Published_Applications_NA_New). Searches run against Amino Acid Published Applications produce two sets of results, with the extensions .rapbm (Published_Applications_AA_Main) and .rapbn (Published_Applications_AA_New).

THIS PAGE IS BLANK

Copyright (c) 1993 - 2006 Compugen Ltd.

OM_nucleic - nucleic search, using sw model

Run on: January 2, 2006, 13:49:34; Search time 308 Seconds
 432.772 Million cell updates/sec

Title: US-09-675-650-4

Perfect score: 20 Gapext 1.0

Sequence: 1 gagttagaaaggatagaacg 20

Scoring table: IDENTITY_NUC

Searched: 4996997 seqs, 333346308 residues

Total number of hits satisfying chosen parameters:

9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21:*

- 1: Geneseqm1980s:*
- 2: Geneseqm1990s:*
- 3: Geneseqm2000s:*
- 4: Geneseqm201as:*
- 5: Geneseqm001bs:*
- 6: Geneseqm002as:*
- 7: Geneseqm2002bs:*
- 8: Geneseqm2003as:*
- 9: Geneseqm2003bs:*
- 10: Geneseqm2003cs:*
- 11: Geneseqm2003ds:*
- 12: Geneseqm2004as:*
- 13: Geneseqm2004bs:*
- 14: Geneseqm2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
c 1	20	100.0	20	4	AACF30668	Aaf30668 Prostate	Aaf30668 Prostate
c 2	20	100.0	506	4	ADFR27554	Aaf30666 Human dif	Aaf30666 Human dif
c 3	20	100.0	506	13	ACN44998	Adr27854 Human pro	Adr27854 Human pro
c 4	17.4	87.0	32323	11	ABZ12053	Acr44098 Human Gen	Acr44098 Human Gen
c 5	16.8	84.0	789	10	ABZ12053	Abz10002 N. gonorr	Abz10002 N. gonorr
c 6	16.8	84.0	792	14	AEB49028	Aaz1053 Neisseria	Aaz1053 Neisseria
c 7	16.8	84.0	792	14	AEB49028	Aeb9028 N. gonorr	Aeb9028 N. gonorr
c 8	16.8	84.0	2086	8	ACRA1244	Aca1244 Prokaryot	Aca1244 Prokaryot
c 9	16.8	84.0	6147	3	AAZ53858	Aaz53858 Neisseria	Aaz53858 Neisseria
c 10	16.8	84.0	35871	4	AAK84978	Aak84974 Human imm	Aak84974 Human imm
c 11	16.4	82.0	65	4	ABN5158	Abn5158 Mouse spl	Abn5158 Mouse spl
c 12	16.4	82.0	102	12	ACR93983	Ach93983 Human gen	Ach93983 Human gen
c 13	16.4	82.0	316	6	ABN23941	Abn23941 Human ORF	Abn23941 Human ORF
c 14	16.4	82.0	347	3	AC119786	Aaz13970 Human gen	Aaz13970 Human gen
c 15	16.4	82.0	540	12	ACH8083	Ach80283 Human gen	Ach80283 Human gen
c 16	15.8	79.0	214	12	ACH92461	Ach2461 Human gen	Ach2461 Human gen
c 17	15.8	79.0	300	2	ABV59234	Abv59234 Human pro	Abv59234 Human pro
c 18	15.8	79.0	304	5	AAK57552	Aak57552 Human imm	Aak57552 Human imm
c 19	15.8	79.0	309	4			

ALIGNMENTS

RESULT 1	
ID	AAF0668 standard; cDNA; 20 BP.
XX	
AC	AAF0668;
XX	
DT	11-JUN-2001 (first entry)
XX	
DE	Prostate cancer antigen 3 (PSA3) nucleic acid.
XX	
KW	PCA3; prostate cancer; antigen; marker; differential expression; diagnosis; therapy; human; ss.
KW	
XX	
OS	Homo sapiens.
XX	
PN	WO200123550-A2.
XX	
PD	05-APR-2001.
XX	
PF	29-SEP-2000; 2000WO-CA001154.
XX	
PR	29-SEP-1999; 99US-0156594P.
XX	
(DIAG-)	DIAGNOURE INC.
PA	
PI	
XX	
Busse U,	Chyrene C, Fradet Y;
DR	WPI; 2001-258132/26.
XX	
PT	The present sequence comprises nucleotides from human prostate cancer antigen 3 (PCA3) long mRNA (see AAF30668). Claimed isolated nucleic acid molecules consist of 10-50 nucleotides which specifically hybridise to a differentially expressed long PCA3 mRNA, and are complementary to, or consist of, at least 10 consecutive nucleotides of the present sequence.
CC	Long PCA3 mRNA contains additional sequence giving rise to long PCA3 mRNA, useful for diagnosis of mammal afflicted with prostate cancer.
PT	
PT	
XX	
PS	Claim 6; Page 57; 60pp; English.
XX	

CC prostate cancer. Long PCA3 RNA is associated with a non-malignant prostatic state. Differential expression of these 2 PCA3 RNA species provides protocols for the diagnosis of prostate disease, including a method of diagnosing the presence or predisposition to develop prostate cancer in a patient. Also provided are therapeutic methods that use a nucleic acid encoding a differentially expressed PCA3 mRNA molecule, an antisense sequence, a protein encoded by a differentially expressed PCA3 mRNA, or an antibody raised against such a protein

SQ Sequence 20 BP; 9 A; 1 C; 8 G; 2 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 18;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTAGGAGGTAGAAACG 20

Db 1 GAGTAGGAGGTAGAAACG 20

RESULT 2

AAF30666/c
ID AAF30666 standard; cDNA; 506 BP.

XX AC AAF30666;

XX DT 11-JUN-2001 (first entry)

XX DE Human differentially expressed PCA3 cDNA (long form).

XX KW PCA3; prostate cancer; antigen; benign prostatic hyperplasia;

KW differential expression; diagnosis; gene therapy; chromosome 9; human;

KW 9B.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT exon 1. .26

XX FT /*tag= a
XX FT /number= 3

XX FT misc_RNA 27. .254

XX FT /*tag= b
XX FT /note= "228 bp insertion"

XX FT exon 255. .506

XX FT /*tag= c
XX FT /number= 4a

XX PN WO200123550-A2.

XX PD 05-APR-2001.

XX PP 29-SEP-2000; 2000WO-CA001154.

XX PR 29-SEP-1999; 99US-0156594P.

XX PA (DIAG-) DIAGNOCURE INC.

XX Busse U, Chypre C, Pradet Y;

XX WPI; 2001-258132/26.

XX Novel nucleic acid encoding differentially expressed prostate cancer

PT antigen 3 mRNA containing additional sequence giving rise to long PCA3

PS mRNA, useful for diagnosis of mammal afflicted with prostate cancer.

XX Claim 3(a); Fig 3; 60pp; English.

XX The present sequence is that of an RT-PCR-amplified fragment of human

PT prostate cancer antigen 3 (PCA3) mRNA that includes an additional 228

PS sequence, inserted between exons 3 and 4a. A 2nd amplified fragment (see

CC AA30667) lacks this additional sequence. The additional sequence

CC interrupts the open reading frame of PCA3 protein, thereby yielding a

CC truncated PCA3 protein. The shorter form PCA3 RNA is associated with

CC prostate cancer whereas the longer form PCA3 RNA is associated with a non
CC -malignant prostatic state, such as benign prostatic hyperplasia. Based
CC on the differential expression of these 2 PCA3 RNA species, protocols for
CC the diagnosis of prostate disease are provided, including a method of
CC diagnosing the presence or predisposition to develop prostate cancer in a
CC patient. Also provided are therapeutic methods that use a nucleic acid
CC encoding a differentially expressed PCA3 mRNA molecule, an antisense
CC sequence, a protein encoded by a differentially expressed PCA3 mRNA, or
CC an antibody raised against such a protein

XX SQ Sequence 506 BP; 132 A; 123 C; 108 G; 143 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 20; DB 4; Length 506;

XX Best Local Similarity 100.0%; Pred. No. 21;

XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACTGGAGGTAGAAACG 20

DB 111 GACTGGAGGTAGAAACG 92

RESULT 3

ADR27854/C

ID ADR27854 standard; DNA; 506 BP.

XX XX AC ADR27854;

XX DT 04-NOV-2004 (first entry)

XX DB Human prostate specific PCA3 sequence long fragment.

XX KW Prostate cancer; PCA3; ds.

XX OS Homo sapiens.

XX PN WO2004070056-A2.

XX PD 19-AUG-2004.

XX PR 09-FEB-2004; 2004WO-CR000170.

XX PR 07-FEB-2003; 2003US-0445436P.

XX PA (DIAG-) DIAGNOCURE INC.

XX PI Pradet Y, Chypre C, Piche L, Garon G;

XX DR WPI; 2004-615607/59.

XX Detecting prostate cancer comprises performing an in vitro nucleic acid amplification assay on a sample using primers specific to PCA3 sequence or prostate specific nucleic acid sequence.

XX Disclosure; SEQ ID NO 7; 97pp; English.

XX The invention relates to a novel method for detecting prostate cancer in a human patient. The method comprises: performing an in vitro nucleic acid amplification assay on a biological sample of the patient or its extract, using a first primer pair, which is specific to a prostate cancer specific PCA3 sequence and a second primer pair, which is specific to a prostate specific nucleic acid sequence, and detecting the PCA3 sequence and the prostate specific nucleic acid sequence, where a detection of the PCA3 nucleic acid sequence or its level correlates with a risk of developing prostate cancer or to a presence of prostate cancer in the patient, and where an absence of detection of the PCA3 nucleic acid sequence or its lower level in the sample validates an absence of prostate cancer or a lower risk of developing prostate cancer, when the second prostate specific nucleic acid is detected. The invention further comprises: a kit for assessing the presence of prostate cancer or the risk of developing prostate cancer in a patient. The method and kits are useful for detecting prostate cancer in a human patient and for determining a predisposition, or presence of prostate cancer or monitoring a progression of prostate cancer in a patient. This

CC polynucleotide sequence represents a fragment of the prostate specific
CC PCA3 gene sequence of the invention.

SQ Sequence 506 BP; 132 A; 123 C; 108 G; 143 T; 0 U; 0 Other;
XX Query Match 100.0%; Score 20; DB 13; Length 506;
CC Best Local Similarity 100.0%; Pred. No. 21;
CC Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GAGTAGGAGGTAGATAAACG 20
Db 111 GAGTAGGAGGTAGATAAACG 92

RESULT 4

ACN41098/C

ID ACN41098 standard; DNA; 32323 BP.

XX AC ACN41098;

XX DT 18-NOV-2004 (first entry)

XX DB Human genomic sequence hCG1641631.

XX KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

XX OS Homo sapiens.

XX PN WO2003073826-A2.

XX PR 01-MAR-2002; 2002US-00087192.

XX PA (SAGR-) SAGRES DISCOVERY.

XX PI Morris DW;

XX PD 12-SEP-2003.

XX PP 28-FEB-2003; 2003WO-US0006235.

XX DR WPI: 2003-328604/31.

XX PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma

XX PT comprises a nucleotide sequence.

XX PS Claim 1; SEQ ID NO 376; Opp; English.

XX CC The present invention relates to novel DNA and protein sequences which

CC are associated with carcinomas. The sequences are useful for:

CC (i) for screening drug candidates;

CC (ii) for detecting Protein (CAP);

CC (iii) for screening of bioactive agent capable

CC of binding to Carcinoma Associated Protein (CAP);

CC (iv) for modulating the activity of CAP;

CC (v) for evaluating the effect of a candidate carcinoma drug;

CC (vi) for inhibiting the activity of CAP;

CC (vii) for neutralizing the effect of CAP;

CC (viii) as a biochip;

CC (ix) for diagnosing carcinoma or a (x) gene copy number. In addition, the

CC genes are useful as DNA vaccines and the CAP are useful as markers of

CC carcinoma including lymphoma. The present sequence is one such CA coding

CC sequence. Note: This patent is an equivalent to basic patent

CC US2002182566A1, for which no sequence data was published

XX SQ Sequence 32323 BP; 7372 A; 7775 C; 8641 G; 8364 T; 0 U; 171 Other;

XX Query Match 87.0%; Score 17.4; DB 11; Length 32323;

CC Best Local Similarity 94.7%; Pred. No. 3 8e+02;

CC Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Qy 1 GAGTAGGAGGTAGATAAAC 19

Db 7698 GAGTAAGGAGGTAGATAAAC 7680

RESULT 5
ABZ40002/C

ID ABZ40002 standard; DNA; 789 BP.

XX AC ABZ40002;

XX DT 07-MAR-2003 (first entry)

XX DE N. gonorrhoeae nucleotide sequence SEQ ID 4593.

XX KW Antibacterial; infection; vaccine; gene therapy; gene; ds.

XX OS Neisseria gonorrhoeae.

XX PN WO200279243-A2.

XX PR 10-OCT-2002.

XX PP 12-FEB-2002; 2002WO-IB002069.

XX PR 12-FEB-2001; 2001GB-00003424.

XX PA (CHIR-) CHIRON SPA.

XX PI Fontana MR, Pizza M, Massignani V, Monaci E;

XX DR WPI: 2003-058415/05.

DR P-PDB; ABP79332.

XX PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a

CC medicament for treating or preventing N. gonorrhoeae infection.

XX PS Disclosure; Page 525; 815pp; English.

XX DR PN

CC The present invention relates to proteins from Neisseria gonorrhoeae.

CC Also disclosed are the nucleic acid molecules encoding the proteins and

CC antibodies that specifically bind to the proteins. The composition

CC comprising the protein, nucleic acid or antibody is useful for the

CC manufacture of a medicament for treating or preventing N. gonorrhoeae

CC infection, this may be in the form of a vaccine or gene therapy

CC sequences given in records ABZ37706-ABZ42016 represent nucleic acid

CC molecules of the invention

XX SQ sequence 789 BP; 208 A; 184 C; 189 G; 208 T; 0 U; 0 Other;

XX Query Match 84.0%; Score 16.8; DB 10; Length 789;

CC Best Local Similarity 90.0%; Pred. No. 5.9e+02;

CC Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX Qy 1 GAGTAGGAGGTAGATAAACG 20

Db 148 GAGTAGGAGGTAGATAAACG 129

XX RESULT 6
AAZ12053/C

ID AAZ12053 standard; DNA; 792 BP.

XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;

CC Treatment; Neisseria; infection; meningitis; septicemia; gonorrhea; ss.

XX AC AAZ12053;

XX DT 08-OCT-1999 (first entry)

XX OS Neisseria gonorrhoeae.

XX PN WO924578-A2.

XX PD 20-MAY-1999.

XX XX

PF 09-OCT-1998; 98W0-IB001665.
 XX
 PR 06-NOV-1997; 97GB-00023516.
 PR 14-NOV-1997; 97GB-00024190.
 PR 97GB-00024386.
 PR 27-NOV-1997; 97GB-00025158.
 PR 10-DEC-1997; 97GB-00026147.
 PR 14-JAN-1998; 98GB-00000759.
 PR 01-SEP-1998; 98GB-00019016.
 XX
 PA (CHIR-) CHIRON SPA.
 PI Masignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;
 XX
 DR WPI: 1999-327407/27.
 DR -PSDB; AAY38591.
 PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for diagnosis, treatment and prevention of infection.
 PT
 XX
 PS Claim 9: Page 150-151; 524pp; English.
 XX
 CC Nucleotide sequences AAY211972-212358 represent open reading frames (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode antigenic proteins (see AAY38499-Y3894). The antigenic proteins, their fragments, their nucleic acids and antibodies are used for diagnosis, prevention (as vaccines) or treatment of Neisseria infections such as meningitis, septicemia and gonorrhea. Both organisms are closely related. Fragments of the nucleic acids are useful as hybridisation probes and antisense reagents.
 CC
 XX
 Sequence 792 BP; 210 A; 184 C; 189 G; 209 T; 0 U; 0 Other.
 Query Match 84.0%; Score 16.8; DB 2; Length 792;
 Best Local Similarity 90.0%; Pred. No. 5.9e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Ps
 Qy 1 GAGTAGGAACGATAGAAACG 20
 Db 148 GAGTAGGAATGAAAGAACG 129
 RESULT 8
 ID ACR41244/C standard; DNA; 2086 BP.
 AC ACR41244;
 XX
 AC ACR41244;
 XX
 DT 27-OCT-2003 (revised)
 XX
 DE Prokaryotic essential gene #22901.
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX
 OS Neisseria gonorrhoeae.
 XX
 WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948933.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362693P.
 XX
 PA (ELITR-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Hasselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trwick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14; SEQ ID NO 29114; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 XX
 PI Scarlato V, Masignani V, Rappuoli R, Pizza M, Grandi G;
 XX
 DR WPI; 2005-464877/47.
 DR -PSDB; AEB49029.

xx	Sequence	2086 BP;	643 A;	461 C;	493 G;	489 T;	0 U;	0 Other;	
SQ	Query Match	84.0%		Score	16.8;	DB	8;	Length	208
	Best Local Similarity	90.0%		Pred. No.	6.2s+02;				
	Matches	18;	Conservative	0;	Mismatches	2;		Indels	
Qy	1	GAGPAGGAACGATGAAACCG	20						
DB	148	GAGPAGGAACGATGAAACCG	129						

RESULT 9
 AA253858/c
 ID AA253858 standard; DNA; 6147 BP.
 XX
 AA253858;
 AC
 XX DT 15-SEP-2003 (revised)
 XX DT 21-MAR-2000 (First entry)
 DE Neisseria gonorrhoeae ORF 564 partial DNA sequence SEQ ID NO:1665.
 XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diognostic; immunogenic; meningitis; infection; meningitis; septicemia;
 KW antibacterial; gene therapy; ds.
 XX OS Neisseria gonorrhoeae.

卷之三

PA	XX	(GENO-) INST GENOMIC RES	
XX		Fraser C, Galeotti C, Gr	
PI		Pettersen J, Pizza M, Rata	
PI		Tectelin H, Venter JC,	
PI	XX	WPI; 2000-062150/05.	
DR	DR	P-PRDB; AAY75096.	
DR	XX	Novel <i>Neisseria</i> polypepti	
PT	PT	vaccines and diagnostics.	
PT	XX	Claim 7; Page 849-851; 14	
PS	XX	AAZ53015 to AAZ54536, AAZ	
CC	CC	represent novel <i>Neisseria</i>	
CC	CC	and polypeptides. AAZ54537-	
CC	CC	PCR primers used in the ex-	
CC	CC	polypeptides, the polynucle-	
CC	CC	invention can be used as v-	
CC	CC	immunogenic compositions.	
CC	CC	of medicaments for treatin-	
CC	CC	bacteria (e.g. meningitis	
CC	CC	<i>Neisseria</i> bacteria, or to	
CC	CC	screen for agonists or anti-	
CC	CC	antibacterial agents. The	
CC	CC	used in gene therapy proto-	
CC	CC	field)	

AA	Sequence 6147 BP; 1998 A; 1476 C; 1444 G; 1229 T; 0 U; 0 Other;
Qy	Query Match 84.0%; Score 16.8; DB 3; Length 6147; Best Local Similarity 90.0%; Pred. No. 6.5e+02; Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Ds	- 1 GAGTAGGAAGGATAGAACG 20 - 148 GAGTAGGAAGGATAGAACG 129

RESULT 10
 AAK84974/C standard; DNA; 35871 BP.
 ID AAK84974
 XX
 AC AAK84974;
 XX
 XX
 DT 07-NOV-2001 (first entry)
 XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:39786.
 DE Human immune/haematopoietic antigen; immune/haematopoietic antigen; cancer;
 KW haemopoietic; immune; haematopoietic; vaccine; metastasis; ds.
 KW cyostatic; gene therapy;
 XX Homo sapiens.
 OS WO200157182-A2.
 PN

PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217495P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225368P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 23-AUG-2000; 2000US-0227758P.
 PR 14-AUG-2000; 2000US-0228266P.
 PR 18-AUG-2000; 2000US-0228267P.
 PR 22-AUG-2000; 2000US-0228681P.
 PR 22-AUG-2000; 2000US-0228688P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 30-AUG-2000; 2000US-0228709P.
 PR 01-SEP-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0228228P.
 PR 01-SEP-2000; 2000US-0228343P.
 PR 01-SEP-2000; 2000US-0228344P.
 PR 01-SEP-2000; 2000US-0228345P.
 PR 05-SEP-2000; 2000US-0228509P.
 PR 08-SEP-2000; 2000US-02282513P.
 PR 08-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 14-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-02323398P.
 PR 14-SEP-2000; 2000US-02323399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0231063P.
 PR 14-SEP-2000; 2000US-0231064P.
 PR 14-SEP-2000; 2000US-0231065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235434P.
 PR 27-SEP-2000; 2000US-0235434P.
 PR 29-SEP-2000; 2000US-0235636P.
 PR 29-SEP-2000; 2000US-0235637P.
 PR 02-OCT-2000; 2000US-0236327P.
 PR 02-OCT-2000; 2000US-0236367P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236374P.
 PR 13-OCT-2000; 2000US-0239355P.
 PR 13-OCT-2000; 2000US-0239337P.
 PR 20-OCT-2000; 2000US-0240560P.
 PR 20-OCT-2000; 2000US-0241211P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0244536P.
 PR 01-NOV-2000; 2000US-0244517P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246529P.
 PR 08-NOV-2000; 2000US-0246530P.
 PR 08-NOV-2000; 2000US-0246535P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246529P.
 PR 08-NOV-2000; 2000US-0246530P.
 PR 08-NOV-2000; 2000US-024610P.
 PR 08-NOV-2000; 2000US-024611P.
 PR 08-NOV-2000; 2000US-024613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249219P.
 PR 17-NOV-2000; 2000US-0249220P.
 PR 17-NOV-2000; 2000US-0249221P.
 PR 17-NOV-2000; 2000US-0249222P.
 PR 17-NOV-2000; 2000US-0249223P.
 PR 17-NOV-2000; 2000US-0249224P.
 PR 17-NOV-2000; 2000US-0249225P.
 PR 17-NOV-2000; 2000US-0249226P.
 PR 17-NOV-2000; 2000US-0249227P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251034P.
 PR 05-DEC-2000; 2000US-0251038P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251860P.
 PR 08-DEC-2000; 2000US-0251860P.
 PR 08-DEC-2000; 2000US-0251865P.
 PR 08-DEC-2000; 2000US-0251865P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Rosen CA, Barash SC, Ruben SM;
 XX DR WPI; 2001-483426/52.
 XX PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX Disclosure; SEQ ID NO 39786; 3071pp + Sequence Listing; English.
 XX PS
 CC AA54951 to AA564702 encode the human immune/hematopoietic antigen (I) amino acid sequences given in AAM82170 to AAM91921. (I) have cytotoxic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the

CC nucleic acids into a host cell and culturing the cell to express the CC protein. (I) proteins and polynucleotides may be used to prevent, CC diagnose and treat immune/haematopoietic diseases, especially CC cancers and cancer metastases of haematopoietic-derived cells. AAK6703 CC to AAK8764 represent human immune/haematopoietic antigen genomic CC sequences from the present invention. AAK5442 to AAK5450 and AAM82169 CC represent sequences used in the exemplification of the present invention XX

SQ Sequence 35871 BP; 10049 A; 6936 C; 7846 G; 11040 T; 0 U; 0 Other;

Query Match 84.0%; Score 16.8; DB 4; Length 35871;

Best Local Similarity 90.0%; Pred. No. 7.1e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGTGGAAAGGATAGAACCG 20

Db 16159 GAGTAGAAAGATAGAACCG 16140

RESULT 11

ABN55158

ID ABN55158 standard; DNA; 65 BP.

XX AC ABN55158;

XX DT 15-JUL-2002 (first entry)

XX DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:27906.

XX PR 20-JUL-2001; 2001W0-IB001903.

XX PR 28-JUL-2000; 2000US-0221607P.

XX PR 02-MAY-2001; 2001US-0287724P.

XX PA (COMP-) COMPUGEN INC.

XX PI Shoshan A, Wasserman A, Mintz E, Mintz L, Raigler S;

XX DR; 2002-257383/30.

XX PS Example 1; SEQ ID NO 27906; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting messenger RNAs that populate a (sub-)transcriptome, where the (sub-) transcriptome comprises messenger RNAs transcribed from multiple transcription units that populate a genome. The library comprises several oligonucleotides, each capable of hybridising selectively to a set of messenger RNAs transcribed from a given transcription unit of the genome, which encodes one or more messenger RNA splice variants. The oligonucleotide libraries are useful for detecting mRNAs from a biological sampler, in expression profiling studies, in qualitatively or quantitatively characterising the corresponding transcriptome, and in detecting RNA transcripts and splice variants of human or animal transcriptomes. The libraries may also be used as specialised mini libraries to detect transcripts of a sub-transcriptome under a particular biological or pathological state, and so allowing the detection of tissue - and pathology specific genes such as those genes only expressed in specific tissue under a specific pathological condition; to detect developmental specific genes; and to detect RNA transcripts and splice

CC variants of a transcriptome of a patient suffering from a particular CC disorder. ABN27253 to ABN59389 represent oligonucleotide sequences from CC humans and mice, which are used in the exemplification of the CC present invention. N.B. The sequence data for this patient did not form CC part of the printed specification, but was obtained in electronic format CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences XX

SQ Sequence 65 BP; 26 A; 12 C; 20 G; 7 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 6; Length 65;

Best Local Similarity 94.4%; Pred. No. 8e-02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGTAGGAAGGATAGAAC 19

Db 19 AGTGGAAAGGATAGAAC 36

RESULT 12

ACH93983^C

ID ACH93983 Standard; DNA; 102 BP.

XX AC ACH93983;

XX DT 29-JUL-2004 (first entry)

XX DE Human genome derived single exon probe #27178.

XX KW Human; probe; ss; gene expression; single exon probe; microarray;

XX KW alternative splicing event; genomic alteration.

XX OS Homo sapiens.

XX PN US200319704-A1.

XX PD 16-OCT-2003.

XX PR 03-APR-2002; 2002US-00029386.

XX PR 03-APR-2002; 2002US-00029386.

XX PA (PENN/) PENN S G.

PA (RANK/) RANK D R.

PA (HANZEL/) HANZEL D K.

XX PI Penn SG, Rank DR, Hanzel DK;

XX DR; 2004-119264/12.

XX WPI; 2004-119264/12.

XX New human genome-derived single exon nucleic acid probes useful for human gene expression analysis, for identifying or characterizing alternative PT splicing events, or as tools for assessing genomic alterations or as tools for PT surveying tissues.

XX Claim 1; SEQ ID NO 27178; 80pp; English.

XX The invention relates to a nucleic acid probe for measuring human gene CC expression, comprising any of the 27,400 fully defined nucleotide CC sequences or their complements or fragments, and CC encoding at least 8 amino acids of any of the 6388 amino acid sequences CC fully defined in the specification. The probe is a single exon probe that CC hybridises under high stringency conditions to a nucleic acid molecule CC expressed in human cells or tissues. Also included are a spatially- CC addressable set of single exon nucleic acid probes for measuring human gene CC expression (comprising a plurality of single exon nucleic acid CC probes cited above, where each of the plurality of probes is separately CC and addressably isolatable or amplifiable from the plurality), a single CC exon microarray for measuring human gene expression, a method of CC measuring human gene expression, a vector comprising the single exon CC probe cited above, an ORF-encoded peptide comprising at least 8 CC contiguous amino acids of any of the above-mentioned amino acid CC sequences (optionally with conservative amino acid substitutions), an CC isolated antibody that binds specifically to a peptide cited above,

methods of selling and/or licensing single exon probes or microarrays to a customer desiring to measure gene expression, a method of providing human gene expression data by subscription, and a computer-readable storage medium which contains a database having a plurality of records (each record including data on the expression of a single exon probe cited above). The probe, methods and apparatus are useful in gene expression analysis. The probes may be used as tools for surveying tissues to detect the presence of expressed messages that contain their specific exon, or in constructing genome-derived single exon microarrays. In addition, the probes used in identifying and characterising gross alternative splicing events, in detecting and characterising gross alterations in the genomic locus that includes their exon, in assessing smaller genomic alterations, in priming the synthesis of nucleic acids, or in expressing the ORF-encoded peptide. The present sequence is a human single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/Sequence.html?DocID=200301194704

CC identified within the present sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3'-untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors.

XX Sequence 347 BP; 132 A; 51 C; 57 G; 105 T; 0 U; 2 Other;

Query Match Score 16.4%; DB 3; Length 347;
Best Local Similarity 94.4%; Pred. No. 8.6e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 AGTAGGAAGGATAGAAC 19
Db 246 AGTAGGAAGGACGAAAC 263

RESULT 15

ID ACH80283 standard; DNA: 540 BP.

XX ACH80283;

XX DT 29-JUL-2004 (First entry)

XX Human genome derived single exon probe #13478.

XX KW Probe; ss; gene expression; single exon probe; microarray;

XX KW alternative splicing event; genomic alteration.

XX OS Homo sapiens.

XX US2003194704-A1.

XX 16-OCT-2003.

XX PF 03-APR-2002; 2002US-00029386.

XX PR 03-APR-2002; 2002US-00029386.

XX PA (PENN/) PENN S G.

PA (RANK/) RANK D R.

PA (HANZL/) HANZEL D K.

XX PI Penn SG, Rank DR, Hanzel DK;

XX DR; 2004-119264/12.

XX New human genome-derived single exon nucleic acid probes useful for human gene expression analysis, for identifying or characterizing alternative splicing events, for assessing genomic alterations or as tools for surveying tissues.

XX PS Claim 15; SEQ ID NO 13478; 80PP; English.

XX DR; 2004-119264/12.

XX The invention relates to a nucleic acid probe for measuring human gene expression, comprising any of the 27,400 fully defined nucleotide sequences in the specification, or their complements or fragments, and encoding at least 8 amino acids of any of the 6888 amino acid sequences fully defined in the specification. The probe is a single exon probe that hybridises under high stringency conditions to a nucleic acid molecule expressed in human cells or tissues. Also included are a spatially-addressable set of single exon nucleic acid probes for measuring human gene expression (comprising a plurality of single exon nucleic acid probes cited above, where each of the plurality of probes is separately and addressably isolatable or amplifiable from the plurality), a single

CC exon microarray for measuring human gene expression, a method of measuring human gene expression, a vector comprising the single exon probe cited above, an ORF-encoded peptide comprising at least 8 contiguous amino acids of any of the above-mentioned amino acid sequences (optionally with conservative amino acid substitutions), an isolated antibody that binds specifically to a peptide cited above, methods of selling and/or licensing single exon probes or microarrays to a customer desiring to measure gene expression, a method of providing human gene expression data by subscription, and a computer-readable storage medium which contains a database having a plurality of records (each record including data on the expression of a single exon probe cited above). The probe, methods and apparatus are useful in gene expression analysis. The probes may be used as tools for surveying tissues to detect the presence of expressed messages that contain their specific exon, or in constructing genome-derived single exon microarrays. In addition, the probes are used in identifying and characterising gross alternative splicing events, in detecting and characterising gross alterations in the genomic locus that includes their exon, in assessing smaller genomic alterations, in priming the synthesis of nucleic acids, or in expressing the ORF-encoded peptide. The present sequence is a human single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?docID=20030194704

SQ Sequence 540 BP; 170 A; 110 C; 79 G; 181 T; 0 U; 0 Other;

SQ Query Match Score 16.4%; DB 12; Length 540;

SQ Best Local Similarity 94.4%; Pred. No. 8.8e+02;

SQ Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAC 18

Db 217 GAGAGGAAGGATAGAAC 200

Search completed: January 2, 2006, 14:05:34
Job time : 312 secs

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 2, 2006, 13:57:24 ; Search time 2243 Seconds
(without alignments)
417.183 Million cell updates/sec

Title: US-09-675-650-4
Perfect score: 20
Sequence: 1 gagtagaaggatagaacg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : EST:
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_hnc: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_est7: *
9: gb_gnb1: *
10: gb_gnb2: *
11: gb_gnb3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	19	95.0	839	10 C2978973	C2978973 195767 To
c 2	18.4	92.0	690	10 CL161954	CL161954 1043531
c 3	18.4	92.0	706	10 CW385789	CW385789 f8b5001f0
c 4	18.4	92.0	776	10 CL158398	CL158398 1043471
c 5	18.4	92.0	812	10 CW385790	CW385790 f8b5001f0
c 6	17.4	87.0	230	5 BW856680	BW856680 BW856680
c 7	17.4	87.0	278	5 BW855726	BW855726 BW855726
c 8	17.4	87.0	475	5 BW853561	BW853561 BW853561
c 9	17.4	87.0	488	5 BW850082	BW850082 BW850082
c 10	17.4	87.0	517	5 BW874566	BW874566 BW874566
c 11	17.4	87.0	525	5 BW871732	BW871732 BW871732
c 12	17.4	87.0	533	2 BG849498	BG849498 1024025e1
c 13	17.4	87.0	551	5 BW741437	BW741437 BW741437
c 14	17.4	87.0	551	5 BW881148	BW881148 BW881148
c 15	17.4	87.0	604	5 BW739479	BW739479 BW739479
c 16	17.4	87.0	616	5 BW320697	BW320697 BW320697
c 17	17.4	87.0	658	9 CL104842	CC104842 CSU-K34.1
c 18	17.4	87.0	941	10 CL076406	CL076406 CH216-139
c 19	17	85.0	549	10 CZ439027	CZ439027 OA_Bba009
c 20	16.8	84.0	180	10 BX650144	BX650144 Arabidops
c 21	16.8	84.0	251	6 CA778813	CA778813 MPL3B4 11
c 22	16.8	84.0	341	5 BQ846596	BQ846596 QGA19078.

ALIGNMENTS

RESULT 1	CZ978973/c	LOCUS	195767 Tomato Mb01 BAC linear	GSS 11-AUG-2005
		DEFINITION	clone SL_Mb010052A23 3, genomic survey sequence.	
		ACCESSION	CZ978973	
		VERSION	CZ978973.1	GI:72330618
		KEYWORDS	GSS.	
		SOURCE	Lycopersicon esculentum (Solanum lycopersicum)	
		ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; asterids; Lamiales; Solanaceae; Solanum; Lycopersicon.	
		REFERENCE	1 (bases 1 to 839)	
		AUTHORS	Mueller,L.A., Bussel,R.M., Wang,Y., Tanksley,S.D., Giovannoni,J.J., van Eck,J. and Stack,S.	
		TITLE	BAC end sequencing from three Solanum lycopersicon libraries	
		JOURNAL	Unpublished (2005)	
		COMMENT	Other GSS: 180444	
		CONTACT	Iukas Mueller	
			Tanksley Lab, Dept. of Plant Breeding	
			Cornell University	
			251 Emerson Hall, Ithaca, NY 14853, USA	
			Tel: 607-255-6557,	
			Fax: 607-255-6683	
			Email: sgn-feedback@sgn.cornell.edu	
			Plate: 62 row: A column: 23	
			Seq primer: SP6	
			Class: BAC ends	
			High quality sequence start: 106	
			High quality sequence stop: 359.	
		FEATURES	Location/Qualifiers	
		Source	1..839	
			/organism="Lycopersicon esculentum"	
			/mol_type="genomic DNA"	
			/culEInv="Heinz 1706"	
			/db_xref="SL Mb010052A23"	
			/lab_host="E. coli"	
			/clonelib="Tomato Mb01 BAC Library"	
			/note="vector: pBioBAC11; site_1: Mb01"	

Matches	19;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;	VERSION	CW385789.1	GI:55104233
Qy	1 GAGTAGGAGGATAGAAC 19									KEYWORDS	GSS.	
Db	813 GAGTAGGAGGATAGAAC 795									SOURCE	Sorghum bicolor (sorghum)	
	RESULT 2									ORGANISM	Sorghum bicolor	
	CL161954 /C	LOCUS	CL161954	690 bp	DNA	linear	GSS 06-JAN-2004			Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; clade; Panicoideae; Andropogoneae; Sorghum.	
	DEFINITION	104 353 10806081 114 31828 369 Sorghum methylation-filtered library (LibID: 104) Sorghum bicolor genomic clone 10806081, genomic survey sequence.								REFERENCE	Bedell, J.A., Budiman, M.A., Nunberg, A., Citek, R.W., Robbins, D., Jones, J., Flick, E., Rohlfing, T., Fries, J., Bradford, K., McMenamy, J., Smith, M., Holman, H., Roe, B.A., Wiley, G., Korf, I.F., Rabinowicz, P.D., Lakey, N., McCombie, W.R., Jeddeloh, J.A. and Martienssen, R.A.	
	ACCESSION	CL161954								AUTHORS	Bedell, J.A., Budiman, M.A., Nunberg, A., Citek, R.W., Robbins, D., Jones, J., Flick, E., Rohlfing, T., Fries, J., Bradford, K., McMenamy, J., Smith, M., Holman, H., Roe, B.A., Wiley, G., Korf, I.F., Rabinowicz, P.D., Lakey, N., McCombie, W.R., Jeddeloh, J.A. and Martienssen, R.A.	
	VERSION	CL161954.1								TITLE	Sorghum genome sequencing by methylation filtration	
	KEYWORDS	GSS.								JOURNAL	PLoS Biol. 3 (1), e13 (2005)	
	SOURCE	Sorghum bicolor (sorghum)								PUBLMED	15660154	
	ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD							COMMENT	Contact: Bedell, JA, Orion Genomics, LLC		
	REFERENCE	1 (bases 1 to 690)								ORGANISM	4041 Forest Park Ave, St. Louis, MO 63108, USA	
	AUTHORS	Bedell, J.A., Budiman, M.A., Nunberg, A., Citek, R.W., Robbins, D., Jones, J., Flick, E., Rohlfing, T., Fries, J., Bradford, K., McMenamy, J., Smith, M., Holman, H., Roe, B.A., Wiley, G., Korf, I.F., Rabinowicz, P.D., Lakey, N., McCombie, W.R., Jeddeloh, J.A. and Martienssen, R.A.							COMMENT	Phone: 314 615 6979 Fax: 314 615 5975 Email: jbedell@oriongenomics.com		
	TITLE	Sorghum genome sequencing by methylation filtration								PLATE	fsbb001f069	
	JOURNAL	PLoS Biol. 3 (1), e13 (2005)								COLLECTION	g column: 09	
	PUBLMED	15660154								PRIMER	314 615 6979	
	COMMENT	Contact: Bedell, JA, Orion Genomics, LLC								SEQUENCE	row: 9	
	ORGANISM	4041 Forest Park Ave, St. Louis, MO 63108, USA								LIBID:	fsbb001f069	
	REFERENCE	TeI: 314 615 6979								CLASS	g	
	AUTHORS	Fax: 314 615 5975								QUALITY	forward	
	DEFINITION	Email: jbedell@oriongenomics.com								STOP	High quality sequence stop: 706.	
	ACCESSION	Plate: 314 615 6979								LOCATION/QUALIFIERS	Location/Qualifiers	
	VERSION	Seq primer: M13/pUC Forward								1.	.706	
	KEYWORDS	Class: methylation filtered								/ORGANISM	/organism="Sorghum bicolor"	
	SOURCE	High quality sequence stop: 690.								/MOL_TYPE	/mol_type="genomic DNA"	
	ORGANISM	Location/Qualifiers								/CULTIVAR	/cultivar="ATx623"	
	REFERENCE	1 (.690)								/DB_XREF	/db_xref="ATx623"	
	AUTHORS	note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells. This is a methylation filtered library." 104)"								/CLONE_LIB	/clone_lib="Sorghum methylation filtered library (LibID: 104)"	
	DEFINITION	note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells. This is a methylation filtered library." 104)"										
	ACCESSION	Qy								ORIGIN	Query Match	
	VERSION	Db								Best Local Similarity	92.0%	Score 18.4;
	KEYWORDS									Pred. No.	DB 10;	Length 706;
	SOURCE									LOCUS	CL158398	GSS 06-JAN-2004
	ORGANISM									DEFINITION	104 347 10803514 114 31377 106 Sorghum methylation-filtered library sequence.	
	REFERENCE									ACCESION	CL158398	
	AUTHORS									VERSION	CL158398	
	DEFINITION									KEYWORDS	1 (bases 1 to 776)	
	ACCESSION									SOURCE	Bedell, J.A., Budiman, M.A., Nunberg, A., Citek, R.W., Robbins, D., Jones, J., Flick, E., Rohlfing, T., Fries, J., Bradford, K., McMenamy, J., Smith, M., Holman, H., Roe, B.A., Wiley, G., Korf, I.F., Rabinowicz, P.D., Lakey, N., McCombie, W.R., Jeddeloh, J.A. and Martienssen, R.A.	
	VERSION									ORGANISM	Sorghum bicolor (sorghum)	
	KEYWORDS									REFERENCE	Sorghum bicolor; Andropogoneae; Sorghum.	
	SOURCE									AUTHORS	Bedell, J.A., Budiman, M.A., Nunberg, A., Citek, R.W., Robbins, D., Jones, J., Flick, E., Rohlfing, T., Fries, J., Bradford, K., McMenamy, J., Smith, M., Holman, H., Roe, B.A., Wiley, G., Korf, I.F., Rabinowicz, P.D., Lakey, N., McCombie, W.R., Jeddeloh, J.A. and Martienssen, R.A.	
	ORGANISM									DEFINITION	CW385789	
	REFERENCE									ACCESSION	CW385789	
	DEFINITION									VERSION	01-NOV-2004	
	ACCESSION									KEYWORDS	Febbb001f06909f0 Sorghum methylation filtered library (LibID: 104)	
	VERSION									SOURCE	Sorghum bicolor genomic clone Febbb001f06909, genomic survey sequence.	
	KEYWORDS									ORGANISM	CW385789	
	SOURCE									REFERENCE	Sorghum genome sequencing by methylation filtration	
	ORGANISM									AUTHORS	Orkin, G., Citek, R.W., Robbins, D., Bedell, J.A., Jones, J., Flick, E., Rohlfing, T., Fries, J., Bradford, K., McMenamy, J., Smith, M., Holman, H., Roe, B.A., Wiley, G., Korf, I.F., Rabinowicz, P.D., Lakey, N., McCombie, W.R., Jeddeloh, J.A. and Martienssen, R.A.	
	DEFINITION									DEFINITION	CW385789	
	ACCESSION									ACCESSION	CW385789	

JOURNAL	PLOS Biol. 3 (1), e13 (2005)	FEATURES	Source	Query Match	92.0%; Score 18.4; DB 10; Length 776;	RESULT 5	Query Match	87.0%; Score 17.4; DB 5; Length 230;
PUBLMED	15660154	COMMENT	Contact: Bedell JA	Best Local Similarity	92.0%; Score 18.4; DB 10; Length 776;	LOCUS	Best Local Similarity	94.7%; Pred. No. 1.5e+03;
			Orion Genomics, LLC	Matches 18;	Matches 1;	DEFINITION	Matches 1;	Mismatches 0;
			4041 Forest Park Ave, St. Louis, MO 63108, USA	Conservative	0;	ORGANISM	Indels 0;	Gaps 0;
			Fax: 314 615 5975			ACCESSION		
			Email: jbedell@oriongenomics.com			VERSION		
			Plate: 347 row: 8 column: 10			KEYWORDS		
			Seq primer: M13/pUC Forward			ORGANISM		
			Class: methylation filtered			REFERENCE		
			High quality sequence stop: 776.			AUTHORS		
		Location/Qualifiers				TITLE		
		1. .776				JOURNAL		
		/organism="Sorghum bicolor"				COMMENT		
		/mol_type="Genomic DNA"						
		/cultivar="ATx623"						
		/db_xref="taxon:4558"						
		/clone="1080514"						
		/clone_lib="Sorghum methylation-filtered library (LibID: 104)"						
		/note="Organ: leaf; Vector: pBCSK(-); Site 1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincI-digested pBCSK(-) vector and electroporated into E. coli cells. This is a methylation filtered library."						
		ORIGIN		Query Match	92.0%; Score 18.4; DB 10; Length 812;	RESULT 6	Query Match	87.0%; Score 17.4; DB 5; Length 230;
				Best Local Similarity	95.0%; Pred. No. 6.4e+02;	LOCUS	Best Local Similarity	94.7%; Pred. No. 1.5e+03;
				Matches 19;	Mismatches 0;	DEFINITION	Matches 18;	Mismatches 0;
				Conservative	0;	ORGANISM	Indels 0;	Gaps 0;
						ACCESSION		
						VERSION		
						KEYWORDS		
						ORGANISM		
						REFERENCE		
						AUTHORS		
						TITLE		
						JOURNAL		
						COMMENT		
		Query Match		Query Match	92.0%; Score 18.4; DB 10; Length 812;	RESULT 7	Query Match	87.0%; Score 17.4; DB 5; Length 230;
		Best Local Similarity		Best Local Similarity	95.0%; Pred. No. 6.4e+02;	LOCUS	Best Local Similarity	94.7%; Pred. No. 1.5e+03;
		Matches 19;		Matches 19;	Mismatches 0;	DEFINITION	Matches 18;	Mismatches 0;
		Conservative		Conservative	0;	ORGANISM	Indels 0;	Gaps 0;
						ACCESSION		
						VERSION		
						KEYWORDS		
						ORGANISM		
						REFERENCE		
						AUTHORS		
						TITLE		
						JOURNAL		
						COMMENT		
		Query Match		Query Match	92.0%; Score 18.4; DB 10; Length 812;	RESULT 8	Query Match	87.0%; Score 17.4; DB 5; Length 230;
		Best Local Similarity		Best Local Similarity	95.0%; Pred. No. 6.4e+02;	LOCUS	Best Local Similarity	94.7%; Pred. No. 1.5e+03;
		Matches 19;		Matches 19;	Mismatches 0;	DEFINITION	Matches 18;	Mismatches 0;
		Conservative		Conservative	0;	ORGANISM	Indels 0;	Gaps 0;
						ACCESSION		
						VERSION		
						KEYWORDS		
						ORGANISM		
						REFERENCE		
						AUTHORS		
						TITLE		
						JOURNAL		
						COMMENT		
		Query Match		Query Match	92.0%; Score 18.4; DB 10; Length 812;	RESULT 9	Query Match	87.0%; Score 17.4; DB 5; Length 230;
		Best Local Similarity		Best Local Similarity	95.0%; Pred. No. 6.4e+02;	LOCUS	Best Local Similarity	94.7%; Pred. No. 1.5e+03;
		Matches 19;		Matches 19;	Mismatches 0;	DEFINITION	Matches 18;	Mismatches 0;
		Conservative		Conservative	0;	ORGANISM	Indels 0;	Gaps 0;
						ACCESSION		
						VERSION		
						KEYWORDS		
						ORGANISM		
						REFERENCE		
						AUTHORS		
						TITLE		
						JOURNAL		
						COMMENT		
		Query Match		Query Match	92.0%; Score 18.4; DB 10; Length 812;	RESULT 10	Query Match	87.0%; Score 17.4; DB 5; Length 230;
		Best Local Similarity		Best Local Similarity	95.0%; Pred. No. 6.4e+02;	LOCUS	Best Local Similarity	94.7%; Pred. No. 1.5e+03;
		Matches 19;		Matches 19;	Mismatches 0;	DEFINITION	Matches 18;	Mismatches 0;
		Conservative		Conservative	0;	ORGANISM	Indels 0;	Gaps 0;
						ACCESSION		
						VERSION		
						KEYWORDS		
						ORGANISM		
						REFERENCE		
						AUTHORS		
						TITLE		
						JOURNAL		
						COMMENT		
		Query Match		Query Match	92.0%; Score 18.4; DB 10; Length 812;	RESULT 11	Query Match	87.0%; Score 17.4; DB 5; Length 230;
		Best Local Similarity		Best Local Similarity	95.0%; Pred. No. 6.4e+02;	LOCUS	Best Local Similarity	94.7%; Pred. No. 1.5e+03;
		Matches 19;		Matches 19;	Mismatches 0;	DEFINITION	Matches 18;	Mismatches 0;
		Conservative		Conservative	0;	ORGANISM	Indels 0;	Gaps 0;
						ACCESSION		
						VERSION		
						KEYWORDS		
						ORGANISM		
						REFERENCE		
						AUTHORS		
						TITLE		
						JOURNAL		
						COMMENT		
		Query Match		Query Match	92.0%; Score 18.4; DB 10; Length 812;	RESULT 12	Query Match	87.0%; Score 17.4; DB 5; Length 230;
		Best Local Similarity		Best Local Similarity	95.0%; Pred. No. 6.4e+02;	LOCUS	Best Local Similarity	94.7%; Pred. No. 1.5e+03;
		Matches 19;		Matches 19;	Mismatches 0;	DEFINITION	Matches 18;	Mismatches 0;
		Conservative		Conservative	0;	ORGANISM	Indels 0;	Gaps 0;
						ACCESSION		
						VERSION		
						KEYWORDS		
						ORGANISM		
						REFERENCE		
						AUTHORS		
						TITLE		
						JOURNAL		
						COMMENT		
		Query Match		Query Match	92.0%; Score 18.4; DB 10; Length 812;	RESULT 13	Query Match	87.0%; Score 17.4; DB 5; Length 230;
		Best Local Similarity		Best Local Similarity	95.0%; Pred. No. 6.4e+02;	LOCUS	Best Local Similarity	94.7%; Pred. No. 1.5e+03;
		Matches 19;		Matches 19;	Mismatches 0;	DEFINITION	Matches 18;	Mismatches 0;
		Conservative		Conservative	0;	ORGANISM	Indels 0;	Gaps 0;
						ACCESSION		
						VERSION		
						KEYWORDS		
						ORGANISM		
						REFERENCE		
						AUTHORS		
						TITLE		
						JOURNAL		
						COMMENT		
		Query Match		Query Match	92.0%; Score 18.4; DB 10; Length 812;	RESULT 14	Query Match	87.0%; Score 17.4; DB 5; Length 230;
		Best Local Similarity		Best Local Similarity	95.0%; Pred. No. 6.4e+02;	LOCUS	Best Local Similarity	94.7%; Pred. No. 1.5e+03;
		Matches 19;		Matches 19;	Mismatches 0;	DEFINITION	Matches 18;	Mismatches 0;
		Conservative		Conservative	0;	ORGANISM	Indels 0;	Gaps 0;
						ACCESSION		
						VERSION		
						KEYWORDS		
						ORGANISM		
						REFERENCE		
						AUTHORS		
						TITLE		
						JOURNAL		
						COMMENT		
		Query Match		Query Match	92.0%; Score 18.4; DB 10; Length 812;	RESULT 15	Query Match	87.0%; Score 17.4; DB 5; Length 230;
		Best Local Similarity		Best Local Similarity	95.0%; Pred. No. 6.4e+02;	LOCUS	Best Local Similarity	94.7%; Pred. No. 1.5e+03;
		Matches 19;		Matches 19;	Mismatches 0;	DEFINITION	Matches 18;	Mismatches 0;
		Conservative		Conservative	0;	ORGANISM	Indels 0;	Gaps 0;
						ACCESSION		
						VERSION		
						KEYWORDS		
						ORGANISM		
						REFERENCE		
						AUTHORS		
						TITLE		
						JOURNAL		
						COMMENT		
		Query Match		Query Match	92.0%; Score 18.4; DB 10; Length 812;	RESULT 16	Query Match	87.0%; Score 17.4; DB 5; Length 230;
		Best Local Similarity		Best Local Similarity	95.0%; Pred. No. 6.4e+02;	LOCUS	Best Local Similarity	94.7%; Pred. No. 1.5e+03;
		Matches 19;		Matches 19;	Mismatches 0;	DEFINITION	Matches 18;	Mismatches 0;
		Conservative		Conservative	0;	ORGANISM	Indels 0;	Gaps 0;
						ACCESSION		
						VERSION		
						KEYWORDS		
						ORGANISM		
						REFERENCE		
						AUTHORS		
						TITLE		
						JOURNAL		
						COMMENT		
		Query Match		Query Match	92.0%; Score 18.4; DB 10; Length 812;	RESULT 17	Query Match	87.0%; Score 17.4; DB 5; Length 230;
		Best Local Similarity		Best Local Similarity	95.0%; Pred. No. 6.4e+02;	LOCUS	Best Local Similarity	94.7%; Pred. No. 1.5e+03;
		Matches 19;		Matches 19;	Mismatches 0;	DEFINITION	Matches 18;	Mismatches 0;
		Conservative		Conservative	0;	ORGANISM	Indels 0;	Gaps 0;
						ACCESSION		
						VERSION		
						KEYWORDS		
						ORGANISM		
						REFERENCE		
						AUTHORS		
						TITLE		
						JOURNAL		
						COMMENT		
		Query Match		Query Match	92.0%; Score 18.4; DB 10; Length 812;	RESULT 18	Query Match	87.0%; Score 17.4; DB 5; Length 230;
		Best Local Similarity		Best Local Similarity	95.0%; Pred. No. 6.4e+02;	LOCUS	Best Local Similarity	94.7%; Pred. No. 1.5e+03;
		Matches 19;		Matches 19;	Mismatches 0;	DEFINITION	Matches 18;	Mismatches 0;
		Conservative		Conservative	0;	ORGANISM	Indels 0;	Gaps 0;
						ACCESSION		
						VERSION		
						KEYWORDS		
						ORGANISM		
						REFERENCE		
						AUTHORS		
						TITLE		
						JOURNAL		
						COMMENT		

ACCESSION VERSION KEYWORDS SOURCE ORGANISM	neurula whole animal Branchiostoma floridae cDNA clone bbne048g21 BW855726 EST Branchiostoma floridae (Florida lancelet) Branchiostoma floridae Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma. 1 (bases 1 to 278)	ORIGIN Query Match Score 87.0%; DB 5; Length 475; Best Local Similarity 94.7%; Pred. No. 1.7e+03; Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
REFERENCE AUTHORS TITLE JOURNAL COMMENT	Yu, J., Holland,L.Z., Shin-i,T., Kohara,Y., Satou,Y. and Satoh,N. Unpublished (2005). Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp.	Qy 2 AGTAAAGGATAGAAACG 20 Db 27 AGTAAAGGATAGAAACG 45
FEATURES Source	1 .278 /organism="Branchiostoma floridae" /mol_type="mRNA" /db_xref="txon:7739" /clone="bbne048g21" /issue_type="whole animal" /dev_stage="neurula" /clone lib="Amphioxus Branchiostoma floridae unpublished cDNA library, neurula whole animal"	RESULT 9 LOCUS BW850082 Amphiouxus Branchiostoma floridae unpublished cDNA library, neurula whole animal Branchiostoma floridae CDNA clone bbne038124 DEFINITION mRNA sequence.
FEATURES Source	1 .278 /organism="Branchiostoma floridae" /mol_type="mRNA" /db_xref="txon:7739" /clone="bbne048g21" /issue_type="whole animal" /dev_stage="neurula" /clone lib="Amphioxus Branchiostoma floridae unpublished cDNA library, neurula whole animal"	ACCESSION BW850082 VERSION EST SOURCE Branchiostoma floridae (Florida lancelet) ORGANISM Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma. 1 (bases 1 to 488)
FEATURES Source	1 .278 /organism="Branchiostoma floridae" /mol_type="mRNA" /db_xref="txon:7739" /clone="bbne038124" /issue_type="whole animal" /dev_stage="neurula" /clone lib="Amphioxus Branchiostoma floridae unpublished cDNA library, neurula whole animal"	REFERENCE AUTHORS Yu,J., Holland,L.Z., Shin-i,T., Kohara,Y., Satou,Y. and Satoh,N. TITLE Unpublished (2005) JOURNAL COMMENT Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp. Location/Qualifiers FEATURES Source 1 .488 /organism="Branchiostoma floridae" /mol_type="mRNA" /db_xref="txon:7739" /clone="bbne038124" /issue_type="whole animal" /dev_stage="neurula" /clone lib="Amphioxus Branchiostoma floridae unpublished cDNA library, neurula whole animal"
FEATURES Source	1 .278 /organism="Branchiostoma floridae" /mol_type="mRNA" /db_xref="txon:7739" /clone="bbne038124" /issue_type="whole animal" /dev_stage="neurula" /clone lib="Amphioxus Branchiostoma floridae unpublished cDNA library, neurula whole animal"	ORIGIN Query Match Score 87.0%; DB 5; Length 488; Best Local Similarity 94.7%; Pred. No. 1.6e+03; Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
ACCESSION VERSION KEYWORDS SOURCE ORGANISM	BW853561 LOCUS BW853561 Amphiouxus Branchiostoma floridae unpublished cDNA library, neurula whole animal Branchiostoma floridae CDNA clone bbne042f21 5', mRNA sequence.	Qy 2 AGTAAAGGATAGAAACG 20 Db 11 AGTAAAGGATAGAAACG 29
REFERENCE AUTHORS TITLE JOURNAL COMMENT	1 (bases 1 to 475) Yu, J., Holland,L.Z., Shin-i,T., Kohara,Y., Satou,Y. and Satoh,N. Unpublished (2005). Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp.	RESULT 10 LOCUS BW874566 DEFINITION mRNA linear EST 24-MAY-2005 neurula whole animal Branchiostoma floridae CDNA clone bbne114g01 5', mRNA sequence.
FEATURES Source	1 .475 /organism="Branchiostoma floridae" /mol_type="mRNA" /db_xref="txon:7739" /clone="bbne042f21" /issue_type="whole animal" /dev_stage="neurula"	ACCESSION VERSION SOURCE BW874566 .1 GI:66489243 EST. Branchiostoma floridae (Florida lancelet) ORGANISM Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma. 1 (bases 1 to 517) Yu, J., Holland,L.Z., Shin-i,T., Kohara,Y., Satou,Y. and Satoh,N. Authors

TITLE	Expressed genes in Branchiostoma floridae
JOURNAL	Unpublished (2005)
COMMENT	Contact: Tadashi Shin-i Center for Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan tel: 81-559-81-6836 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp.
FEATURES	Location/Qualifiers
source	/organism="Branchiostoma floridae" 'mol_type'='mRNA' 'db_xref'='taxon:7739' 'clone'='bbne114g01' 'issue_stage'='whole animal' 'dev_stage'='neurula' 'clone_lib'='Amphioxus Branchiostoma floridae unpublished cDNA Library, neurula whole animal'
ORIGIN	
Query Match	Score 17.4; DB 5; Length 517;
Best Local Similarity	94.7%
Matches	18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db	317 AGTAAAGGTAGAACG 20
Db	311 AGTAAAGGTAGAACG 335
RESULT 11	
LOCUS	BW871732
DEFINITION	525 bp mRNA linear EST 24-MAY-2005
ORGANISM	Branchiostoma floridae unpublished cDNA library, neurula whole animal Branchiostoma floridae cDNA clone bbnel06c12 5', mRNA sequence.
REFERENCE	BW871732.1 GI:66486409
AUTHORS	Yu J., Holland L.Z., Shin-i T., Kohara Y., Satoh N.
TITLE	Expressed genes in Branchiostoma floridae
JOURNAL	Unpublished (2005)
COMMENT	Contact: Tadashi Shin-i Center for Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp.
FEATURES	Location/Qualifiers
source	/organism="Branchiostoma floridae" 'mol_type'='mRNA' 'db_xref'='taxon:7739' 'clone'='bbne106c12' 'issue_stage'='whole animal' 'dev_stage'='neurula' 'clone_lib'='Amphioxus Branchiostoma floridae unpublished cDNA Library, neurula whole animal'
ORIGIN	
Query Match	Score 17.4; DB 5; Length 525;
Best Local Similarity	94.7%
Matches	18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db	228 ACTGGAGGTAGAACG 20
Db	228 ACTGGAGGTAGAACG 246
RESULT 12	
LOCUS	BG849498
DEFINITION	533 bp mRNA linear EST 29-MAY-2001
ORGANISM	Chlamydomonas reinhardtii cDNA, mRNA sequence.
REFERENCE	1 (bases 1 to 533) Grossman,A., Davies,J., Federspiel,N., Harris,B., LeFebvre,P., McDermott,J.P., Sillflow,C., Stern,D. and Surycki,R.
AUTHORS	Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants, project phase 2
TITLE	Unpublished (2000)
JOURNAL	Contact: Charles Hauser DCMB Box 91000 Duke University Durham, NC 27708-1000
COMMENT	Tel: 919 613 8159 Fax: 919 613 8177 Email: chauser@duke.edu.
FEATURES	Source Location/Qualifiers
Db	1 533 /organism="Chlamydomonas reinhardtii" /mol_type="mRNA" /strain="CC-1690 wild type mt+ 21gr" /db_xref="taxon:3055" /clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap II" /note="vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XbaI; This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (acetate-containing) medium in the light, TAP medium in the dark, HS (minimal) medium in ambient levels of CO ₂ and HS medium bubbled with 5% CO ₂ . PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into Lambda Zap II (Stratagene) in the EcoRI (5') and XbaI (3') sites. pBluescript II SK plasmids were excised from the lambda ZAP clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996). Genome Research 6: 791-806."
ORIGIN	
Query Match	87.0%; Score 17.4; DB 2; Length 533;
Best Local Similarity	94.7%; Pred. No. 1.8e+03;
Matches	18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db	1 GAGTAGGAGGTAGAAC 19 321 GAGTGAAGGTAGAAC 303
RESULT 13	
LOCUS	BW741437
DEFINITION	551 bp mRNA linear EST 09-AUG-2005
ORGANISM	Amphioxus Branchiostoma floridae unpublished cDNA library, mRNA sequence.
REFERENCE	BW741437 Amphioxus Branchiostoma floridae (Florida lancelet) egg whole animal Branchiostoma floridae cDNA clone bbeg039n13 5', mRNA sequence.
AUTHORS	BW741437.1 GI:66328085
TITLE	Branchiostoma floridae (Florida lancelet)
JOURNAL	Branchiostoma floridae
COMMENT	Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.
FEATURES	
Db	

REFERENCE	1. (bases 1 to 551)	Best Local Similarity 94.7%; Pred. No. 1.8e+03;
AUTHORS	Yu,J., Holland,L.Z., Shin-i,T., Kohara,Y., Satou,Y. and Satoh,N.	Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
JOURNAL	Unpublished (2005)	
COMMENT	Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp If you want to have a cDNA clone for this EST or if you have any questions, please send an e-mail to Nori Satoh (satoh@ascidian.zool.kyoto-u.ac.jp) and its cc to Yutaka Satou (yutaka@ascidian.zool.kyoto-u.ac.jp).	
FEATURES		
source	1. 551 /organism="Branchiostoma floridae" /mol_type="mRNA" /db_xref="txaxon:7739" /clone="bbeg025f075" /tissue_type="whole animal" /dev_stage="egg" /clone_lib="Amphioxus Branchiostoma floridae unpublished cDNA library, egg whole animal"	
ORIGIN		
	Query Match Score 87.0%; Score 17.4; DB 5; Length 551; Best Local Similarity 94.7%; Pred. No. 1.8e+03; Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
Qy	2 AGTAGGAGGTAGAAACG 20	
Db	97 AGTAGGAGGTAGAAACG 115	
RESULT 14		
LOCUS	BW881148	551 bp mRNA linear EST 24-MAY-2005
DEFINITION	BW881148 Amphioxus Branchiostoma floridae unpublished cDNA library, neurula whole animal Branchiostoma floridae cDNA clone bbne095n04	
5', mRNA sequence.		
ACCESSION	BW881148	
VERSION	BB881148.1 GI: 66495825	
KEYWORDS	EST,	
SOURCE	Branchiostoma floridae (Florida lancelet)	
ORGANISM	Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma. Branchiostoma.	
REFERENCE	1. (bases 1 to 551)	
AUTHORS	Yu,J., Holland,L.Z., Shin-i,T., Kohara,Y., Satou,Y. and Satoh,N.	
JOURNAL	Expressed Genes in Branchiostoma floridae Unpublished (2005)	
COMMENT	Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp. Location/Qualifiers	
FEATURES		
source	1. 551 /organism="Branchiostoma floridae" /mol_type="mRNA" /db_xref="txaxon:7739" /clone="bbne095n04" /tissue_type="whole animal" /dev_stage="egg" /clone_lib="Amphioxus Branchiostoma floridae unpublished cDNA library, egg whole animal"	
ORIGIN		
	Query Match Score 87.0%; Score 17.4; DB 5; Length 604;	
Qy	2 AGTAGGAGGTAGAAACG 20	
Db	161 AGTAGGAGGTAGAAACG 179	
FEATURES		
source	1. 551 /organism="Branchiostoma floridae" /mol_type="mRNA" /db_xref="txaxon:7739" /clone="bbne095n04" /tissue_type="whole animal" /dev_stage="neurula" /clone_lib="Amphioxus Branchiostoma floridae unpublished cDNA library, neurula whole animal"	
ORIGIN		
	Query Match Score 87.0%; Score 17.4; DB 5; Length 551;	

Search completed: January 2, 2006, 15:18:40
Job time : 2246 secs

GenCore .version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 2, 2006, 14:00:00 ; Search time 104 Seconds
(without alignments)
341.839 Million cell updates/sec

Title: US-09-675-650-4
Perfect score: 20
Sequence: 1 gagtaggaaggatagaaacg 20

Scoring table: IDENTITY_NUC
Gapop 10-0 , Gapext 1.0

Searched: 1303057 seqs, 88780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/picodata/1/ina/1-COMB.seq:*

2: /cgn2_6/picodata/1/ina/5-COMB.seq:*

3: /cgn2_6/picodata/1/ina/6A-COMB.seq:*

4: /cgn2_6/picodata/1/ina/6B-COMB.seq:*

5: /cgn2_6/picodata/1/ina/PCUTS-COMB.seq:*

6: /cgn2_6/picodata/1/ina/PP-COMB.seq:*

7: /cgn2_6/picodata/1/ina/RE-COMB.seq:*

8: /cgn2_6/picodata/1/ina/batchfile1.seq:*

9: /cgn2_6/picodata/1/ina/batchfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17.4	87.0	601	3 US-09-949-016-40733	Sequence 40733, A
C 2	17.4	87.0	465589	3 US-09-949-016-12900	Sequence 12900, A
C 3	17.4	87.0	476944	3 US-09-949-016-12412	Sequence 12412, A
C 4	16.4	82.0	347	3 US-09-513-992861	Sequence 23861, A
C 5	16.4	82.0	312474	3 US-09-949-016-17334	Sequence 17334, A
C 6	15.8	80.0	117838	3 US-09-949-016-17595	Sequence 17595, A
C 7	15.8	79.0	601	3 US-09-949-016-11749	Sequence 11749, A
C 8	15.8	79.0	601	3 US-09-949-016-16978	Sequence 16978, A
C 9	15.8	79.0	41736	3 US-09-949-016-17091	Sequence 17091, A
C 10	15.8	79.0	61178	3 US-09-949-016-17369	Sequence 17369, A
C 11	15.8	79.0	67386	3 US-09-949-016-16519	Sequence 16519, A
C 12	15.8	79.0	94503	3 US-09-949-016-16324	Sequence 16324, A
C 13	15.8	79.0	111135	3 US-09-949-016-15288	Sequence 15288, A
C 14	15.8	79.0	114942	3 US-09-949-016-14933	Sequence 14933, A
C 15	15.4	77.0	601	3 US-09-949-016-92291	Sequence 92291, A
C 16	15.4	77.0	601	3 US-09-949-016-92292	Sequence 92292, A
C 17	15.4	77.0	601	3 US-09-949-016-92293	Sequence 92293, A
C 18	15.4	77.0	601	3 US-09-949-016-92294	Sequence 92294, A
C 19	15.4	77.0	601	3 US-09-949-016-121635	Sequence 121635, A
C 20	15.4	77.0	2973	3 US-10-1-04-047-634	Sequence 634, APP
C 21	15.4	77.0	38009	3 US-09-949-016-13617	Sequence 13617, A
C 22	15.4	77.0	69373	3 US-10-949-016-15140	Sequence 15140, A
C 23	15.4	77.0	105413	3 US-10-27-923-3	Sequence 3, Appli
C 24	15.4	77.0	112219	3 US-09-949-016-12453	Sequence 12453, A

ALIGNMENTS

RESULT 1
US-09-949-016-40733
; Sequence 40733, Application US/0949016
; Patent No. 6812319
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40733
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-40733
Query Match 87.0%; Score 17.4%; DB 3; Length 601;
Best Local Similarity 94.7%; Pred. No. 37;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGTGGAAAGGATAGAAAC 19
Db 446 GAGTGGAAAGGATAGAAAC 464

RESULT 2
US-09-949-016-12900/c
; Sequence 12900, Application US/0949016
; Patent No. 6812319
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

```

; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 12900
; LENGTH: 462589
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-12900

Query Match Score 17.4%; DB 3; Length 462589;
Best Local Similarity 94.7%; Pred. No. 93;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GAGTAGGAAGGATAGAAC 19
Db 132852 GAGTAGGAAGGAAAGAAC 132834

RESULT 3
US-09-949-016-12412/C
; Sequence 12412, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 12412
; LENGTH: 476044
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-12412

Query Match Score 17.4%; DB 3; Length 476044;
Best Local Similarity 94.7%; Pred. No. 93;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GAGTAGGAAGGATAGAAC 19
Db 146308 GAGTAGGAAGGAAAGAAC 146290

RESULT 4
US-09-513-999C-23861
; Sequence 23861, Application US/09513999C
; Patent No. 6783951
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783951
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO: 23861
; LENGTH: 347
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: misc_feature
; NAME/KEY: misc_feature
; LOCATION: 10
; OTHER INFORMATION: w=a or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 311
; OTHER INFORMATION: r=a or g
; US-09-513-999C-23861

```

```

RESULT 5
US-09-949-016-17434
; Sequence 17434, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 17434
; LENGTH: 312474
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-17434

```

```

Query Match Score 16.4%; DB 3; Length 347;
Best Local Similarity 94.4%; Pred. No. 1 1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 ACTAGGAGGGATAGAAC 19
Db 246 ACTAGGAGGGACGAAAC 263

```

```

RESULT 6
US-09-949-016-17595
; Sequence 17595, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 17595
; LENGTH: 312474
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-17595

```

```

Query Match Score 16.4%; DB 3; Length 312474;
Best Local Similarity 94.4%; Pred. No. 2 7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 ACTAGGAGGGATAGAAC 19
Db 9872 AGTTGAGGGATAGAAC 9889

```

NUMBER OF SEQ ID NOS: 207012 SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 17595 LENGTH: 117838
; OTHER INFORMATION: n = A, T, C or G
; TYPE: DNA
; ORGANISM: Human
; FEATURE: misc_feature
; LOCATION: (1)..(117838)
; OTHER INFORMATION: n = A, T, C or G
; SEQ ID NO: 09-949-016-17595

Query Match Score 80.0%; DB 3; Length 117838;
Best Local Similarity 100.0%; Pred. No. 3.7e-02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TAGGAGGATAGAAC 19
Db 15884 TAGGAGGATAGAAC 15899

RESULT 9
US-09-949-016-17091 Application US/09949016
; Sequence 17091, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 117449
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; SEQ ID NO: 09-949-016-117449

Query Match Score 79.0%; DB 3; Length 601;
Best Local Similarity 89.5%; Pred. No. 2.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGTGGAAAGGTAGAAC 19
Db 498 GAGTGGCTGGTAGAAC 480

RESULT 10
US-09-949-016-17369 Application US/09949016
; Sequence 17369, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 17369
; LENGTH: 61178
; TYPE: DNA
; ORGANISM: Human

RESULT 8
US-09-949-016-169788/C
; Sequence 169788, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012

US-09-949-016-17369

Query Match	79.0%	Score 15.8;	DB 3;	Length 61178;
Best Local Similarity	89.5%;	Pred. No. 4.2e+02;		
Matches 17;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	1 GAGTAGAAGGATAGAAC 19			
Db	44335 GAGTATGCCGGATAGAAC 44353			

RESULT 11

US-09-949-016-16519	Sequence 16519, Application US/09949016
;	Patent No. 6812339
;	GENERAL INFORMATION:
;	APPLICANT: VENTER, J. Craig et al.
;	TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;	CURRENT APPLICATION NUMBER: US/09/949,016
;	FILE REFERENCE: CLO01307
;	CURRENT FILING DATE: 2000-04-14
;	PRIOR APPLICATION NUMBER: 60/241,755
;	PRIOR FILING DATE: 2000-10-20
;	PRIOR APPLICATION NUMBER: 60/237,768
;	PRIOR FILING DATE: 2000-10-03
;	PRIOR APPLICATION NUMBER: 60/231,498
;	PRIOR FILING DATE: 2000-09-08
;	NUMBER OF SEQ ID NOS: 207012
;	SOFTWARE: FastSEQ for Windows Version 4.0
;	SEQ ID NO 16519
;	LENGTH: 67386
;	TYPE: DNA
;	ORGANISM: Human
;	FEATURE:
;	NAME/KEY: misc_feature
;	LOCATION: (1) ..(67386)
;	OTHER INFORMATION: n = A,T,C or G

US-09-949-016-16519

Query Match	79.0%	Score 15.8;	DB 3;	Length 67386;
Best Local Similarity	89.5%;	Pred. No. 4.3e+02;		
Matches 17;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	1 GAGTAGAAGGATAGAAC 19			
Db	3862 GACTGAAGGATAGAAC 3880			

RESULT 12

US-09-949-016-16324/C	Sequence 16324, Application US/09949016
;	Patent No. 6812339
;	GENERAL INFORMATION:
;	APPLICANT: VENTER, J. Craig et al.
;	TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;	CURRENT APPLICATION NUMBER: US/09/949,016
;	FILE REFERENCE: CLO01307
;	CURRENT FILING DATE: 2000-04-14
;	PRIOR APPLICATION NUMBER: 60/241,755
;	PRIOR FILING DATE: 2000-10-20
;	PRIOR APPLICATION NUMBER: 60/237,768
;	PRIOR FILING DATE: 2000-10-03
;	PRIOR APPLICATION NUMBER: 60/231,498
;	PRIOR FILING DATE: 2000-09-08
;	NUMBER OF SEQ ID NOS: 207012
;	SOFTWARE: FastSEQ for Windows Version 4.0
;	SEQ ID NO 16324
;	LENGTH: 94593
;	TYPE: DNA
;	FEATURE:

TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(114842)
OTHER INFORMATION: n = A, T, C or G
US-09-949-016-14993

Query Match 79.0%; Score 15.8; DB 3; Length 114842;
Best Local Similarity 89.5%; Pred. No. 4.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGTAGAAGGATAGAAC 19
Db 45824 GAGTAGGCCTGATGAAAC 45806

RESULT 15
US-09-949-016-92291
Sequence 92291, Application US/09949016
Patent No. 612339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 92291
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-92291

Query Match 77.0%; Score 15.4; DB 3; Length 601;
Best Local Similarity 94.1%; Pred. No. 3.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGTAGAAGGATAGAAC 17
Db 509 GAGTAGAAGGATACAA 525

Search completed: January 2, 2006, 15:20:30
Job time : 107 secs

THIS PAGE IS BLANK

Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 2, 2006, 14:05:41 ; Search time 451 Seconds
(without alignments)
366.713 Million cell updates/sec

Title: US-09-675-650-4

Perfect score: 20

Sequence: 1 gagtaggaaggatagaacg 20

Scoring table: IDENTITY_NUC Gapop 10.0 , Gapext 1.0

Searched: 979542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : Published Applications NA_Main:
1: /cgn2_6/picodata/1/pubnra/US07_PUBCOMB.seq:
2: /cgn2_6/picodata/1/pubnra/US08_PUBCOMB.seq:
3: /cgn2_6/picodata/1/pubnra/US09_PUBCOMB.seq:
4: /cgn2_6/picodata/1/pubnra/US09B_PUBCOMB.seq:
5: /cgn2_6/picodata/1/pubnra/US10A_PUBCOMB.seq:
6: /cgn2_6/picodata/1/pubnra/US10B_PUBCOMB.seq:
7: /cgn2_6/picodata/1/pubnra/US10C_PUBCOMB.seq:
8: /cgn2_6/picodata/1/pubnra/US10D_PUBCOMB.seq:
9: /cgn2_6/picodata/1/pubnra/US10E_PUBCOMB.seq:
10: /cgn2_6/picodata/1/pubnra/US11_PUBCOMB.seq:
*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the core of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	17.4	87.0	348	7 US-10-424-599-12713	Sequence 12713, A
C 2	17.4	87.0	3223	5 US-10-087-19-376	Sequence 376, App
C 3	16.8	84.0	290	8 US-10-125-115-17876	Sequence 17876, App
C 4	16.8	84.0	553	4 US-09-925-065A-267785	Sequence 267785,
C 5	16.8	84.0	2086	7 US-10-285-122A-29114	Sequence 29114, A
C 6	16.8	84.0	9924	6 US-10-085-117-117	Sequence 115, App
C 7	16.4	82.0	65	3 US-09-908-97-27906	Sequence 27906, A
C 8	16.4	82.0	102	6 US-10-059-386-27178	Sequence 27178, A
C 9	16.4	82.0	528	4 US-09-925-065A-512839	Sequence 542839,
C 10	16.4	82.0	540	6 US-10-029-386-134748	Sequence 13478, A
C 11	16.4	82.0	549	4 US-09-925-065A-32759	Sequence 32759,
C 12	16.4	82.0	582	4 US-09-925-065A-63703	Sequence 63703,
C 13	16.4	82.0	600	9 US-10-972-079-67862	Sequence 67862, A
C 14	16.4	82.0	600	9 US-10-972-079-67863	Sequence 67863, A
C 15	16.4	82.0	770	4 US-09-925-065A-70545	Sequence 70545, A
C 16	16.4	82.0	2300	5 US-10-027-632-103227	Sequence 103227,
C 17	16.4	82.0	2300	5 US-10-027-632-103228	Sequence 103228,
C 18	16.4	82.0	2300	6 US-10-027-632-103228	Sequence 103228,
C 19	16.4	82.0	2300	6 US-10-027-632-103228	Sequence 103228,
C 20	16	80.0	412	7 US-10-424-599-133667	Sequence 133667,
C 21	16	80.0	579	4 US-09-925-065A-541153	Sequence 541153,
C 22	16	80.0	579	4 US-09-925-065A-541154	Sequence 541154,
C 23	15.8	79.0	214	6 US-10-029-386-25656	Sequence 25656, A

RESULT 1
US-10-424-599-12713/c
; Sequence 12713, Application US/10424599
; Publication No. US20040031172A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Zhou Yihua
; APPLICANT: Kovacic David K
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 12713
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) .(348)
; OTHER INFORMATION: unsure at all n locations

RESULT 1
US-10-424-599-12713/c
; Sequence 12713, Application US/10424599
; Publication No. US20040031172A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Zhou Yihua
; APPLICANT: Kovacic David K
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 12713
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) .(348)
; OTHER INFORMATION: unsure at all n locations

RESULT 2
US-10-087-192-376/c
; Sequence 376, Application US/10087192
; Publication No. US20020182386A1
; GENERAL INFORMATION:
; APPLICANT: Englehardt, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000122

CURRENT APPLICATION NUMBER: US/10/087,192
 CURRENT FILING DATE: 2002-03-01
 PRIORITY NUMBER: US 09/747,377
 PRIOR FILING DATE: 2000-12-22
 PRIORITY NUMBER: US 09/798,586
 PRIOR FILING DATE: 2001-03-02
 NUMBER OF SEQ ID NOS: 2059
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 3216
 LENGTH: 32323
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1) .. (32323)
 OTHER INFORMATION: n = A,T,C or G
 US-10-087-192-376

Query Match 87.0%; Score 17.4%; DB 5; Length 32323;
 Best Local Similarity 94.7%; Pred. No. 3.1e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GAGTAGGAAGGATAGAAC 19
 Db 7698 GAGTAGGAAGGATAGAAC 7680

RESULT 3

US-10-425-115-178767/c
 Sequence 178167, Application US/10425115
 Publication No. US20040214272A1
 GENERAL INFORMATION:
 APPLICANT: La Rosa, Thomas J.
 APPLICANT: Kovalic, David K.
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With Plants
 FILE REFERENCE: 38-21153222
 CURRENT APPLICATION NUMBER: US/10/425,115
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 369326
 SEQ ID NO: 178767
 LENGTH: 290
 TYPE: DNA
 ORGANISM: Zea mays
 FEATURE:
 NAME/KEY: unsure
 LOCATION: (1) .. (290)
 OTHER INFORMATION: unsure at all n locations
 FEATURE:
 OTHER INFORMATION: Clone ID: MRT4577_94619C.1

Query Match 84.0%; Score 16.8%; DB 8; Length 290;
 Best Local Similarity 90.0%; Pred. No. 4.2e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 GAGTAGGAAGGATAGAAC 20
 Db 58 GAGGGAGGAAGGAAAGAACG 39

RESULT 4

US-09-925-065A-267785/c
 Sequence 267785, Application US/09925065A
 Publication No. US20050228172A9
 GENERAL INFORMATION:
 APPLICANT: Wang, David G.
 TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
 FILE REFERENCE: 108827.135
 CURRENT APPLICATION NUMBER: US/09/25,065A
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 29114

CURRENT FILING DATE: 2001-08-08
 PRIORITY NUMBER: US 60/243,096
 PRIOR FILING DATE: 2000-10-24
 PRIORITY NUMBER: US 60/252,147
 PRIOR FILING DATE: 2000-11-20
 PRIORITY NUMBER: US 60/250,092
 PRIOR FILING DATE: 2000-11-30
 PRIORITY NUMBER: US 60/261,766
 PRIOR FILING DATE: 2001-01-16
 PRIORITY NUMBER: US 60/289,846
 PRIOR FILING DATE: 2001-05-09
 NUMBER OF SEQ ID NOS: 957086
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 267785
 LENGTH: 553
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-925-065A-267785

RESULT 5

US-10-282-122A-29114/c
 Sequence 29114, Application US/10282122A
 Publication No. US200402912911
 GENERAL INFORMATION:
 APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselbeck, Robert
 APPLICANT: Olsen, Kari
 APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ELTRA-034A
 CURRENT APPLICATION NUMBER: US/10/282,122A
 CURRENT FILING DATE: 2003-02-20
 PRIORITY NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIORITY NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIORITY NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIORITY NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIORITY NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIORITY NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIORITY NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIORITY NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIORITY NUMBER: 60/267,536
 PRIOR FILING DATE: 2001-02-09
 PRIORITY NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 29114

LENGTH: 2086
 TYPE: DNA
 ORGANISM: *Neisseria gonorrhoeae*
 US-10-282-122A-29114

Query Match Score 16.8; DB 7; Length 2086;
 Best Local Similarity 90.0%; Pred. No. 4.9e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGTAGGAAGGATAGAAACG 20
 Db 148 GAGTAGGATAGAAAGAACG 129

RESULT 8
 US-10-029-386-27178/c
 Sequence 27178, Application US/10029386
 ; Sequence 27178, Application US/10029386
 ; Publication No. US2003019704A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
 ; FILE REFERENCE: ABOMICA-X-2
 ; CURRENT APPLICATION NUMBER: US/10/029,386
 ; CURRENT FILING DATE: 2001-12-20
 ; NUMBER OF SEQ ID NOS: 34288
 ; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO: 27178
 ; LENGTH: 102
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO CHR14.3
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW. SIGNAL = 0.63
 ; OTHER INFORMATION: SWISSPROT HIT: P20917, EVALUATE 8.00e+00
 ; OTHER INFORMATION: NT HIT: AJ002571.1, EVALUATE 1.30e+00
 US-10-029-386-27178

Query Match Score 16.4; DB 6; Length 102;
 Best Local Similarity 94.4%; Pred. No. 6e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGTAGGAAGGATAGAAAC 19
 Db 19 AGTAGGAAGGATAGAAC 36

RESULT 9
 US-09-925-065A-542039
 Sequence 542039, Application US/09925065A
 ; Sequence 542039, Application US/09925065A
 ; Publication No. US20050228872A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single
 ; Nucleotide Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.125
 ; CURRENT APPLICATION NUMBER: US/09/925,065A
 ; PRIOR APPLICATION NUMBER: 2000-08-08
 ; PRIOR APPLICATION NUMBER: US 60/243,096
 ; PRIOR FILING DATE: 2000-10-24
 ; PRIOR APPLICATION NUMBER: US 60/252,147
 ; PRIOR FILING DATE: 2000-11-20
 ; PRIOR APPLICATION NUMBER: US 60/250,092
 ; PRIOR FILING DATE: 2000-11-30
 ; PRIOR APPLICATION NUMBER: US 60/261,766
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/289,846
 ; PRIOR FILING DATE: 2001-05-09
 ; NUMBER OF SEQ ID NOS: 957086
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 542039
 ; LENGTH: 528
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens

RESULT 7
 US-09-908-975-27906
 Sequence 27906, Application US/09908975
 ; Sequence 27906, Application US/09908975
 ; Publication No. US20030165843A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SHOSHAN, Avi
 ; APPLICANT: WASSERMAN, Alon
 ; APPLICANT: MINTZ, Eli
 ; APPLICANT: FAIGLER, Simchon
 ; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
 ; FILE REFERENCE: 36688-0005
 ; CURRENT APPLICATION NUMBER: US/09/908,975
 ; CURRENT FILING DATE: 2001-07-20
 ; PRIOR APPLICATION NUMBER: US 60/287,724
 ; PRIOR FILING DATE: 2001-05-02
 ; PRIOR APPLICATION NUMBER: US 60/221,607
 ; PRIOR FILING DATE: 2000-07-28
 ; NUMBER OF SEQ ID NOS: 32337
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 27906
 ; LENGTH: 65
 ; TYPE: DNA
 ; ORGANISM: *Mus musculus*

RESULT 10
US-09-925-065A-542839
Query Match Score 16.4; DB 4; Length 528;
Best Local Similarity 82.0%; Pred. No. 6.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 3 GTAGGAAGGATAGAAACG 20
Db 101 GTAGGAAGGATAGAAATG 118

RESULT 11
US-09-925-065A-327859/C
Sequence 327859, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE REFERENCE: AEONICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO: 13478
; LENGTH: 540
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR14.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW SIGNAL = 0.63
; OTHER INFORMATION: NT HIT: UB7252.1, EVALUE 3.30e-02
; OTHER INFORMATION: SWISSPROT HIT: Q03560, EVALUE 2.30e+00
; SEQ ID NO: 10-029-386-13478

Query Match Score 16.4; DB 6; Length 540;
Best Local Similarity 94.4%; Pred. No. 6.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GAGTAGCAAGGATAGAA 18
Db 217 GAGAACGAAAGATGAA 200

RESULT 12
US-09-925-065A-635703
Sequence 635703, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 635703
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Homo sapiens
; SEQ ID NO: 09-925-065A-635703

Query Match Score 16.4; DB 4; Length 582;
Best Local Similarity 94.4%; Pred. No. 6.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 3 GTAGGAAGGATAGAAACG 20
Db 544 GTAGGAAGGATAGAAATG 561

RESULT 13
US-10-972-079-67862/C
Sequence 67862, Application US/10972079
; Publication No. US20050153317A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: ROSENFIELD, David
; APPLICANT: KERR, Richard
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEEF LIVESTOCK
; FILE REFERENCE: MM1110-2
; CURRENT APPLICATION NUMBER: US/10/972,079
; CURRENT FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 60/514,333
; PRIOR FILING DATE: 2003-10-24
; NUMBER OF SEQ ID NOS: 96531
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO: 67862
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Chicken
; SEQ ID NO: 327859
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens

Best Value Copy

STIC-Biotech/ChemLib

175060
Mg

From: Yu, Misook
Sent: Tuesday, December 27, 2005 7:09 AM
To: STIC-Biotech/ChemLib
Subject: 09/675,650

Pls search SEQ ID NO: 4 (20 nts primer)

Examiner Misook Yu, Ph.D.
 571-272-0839 (Phone)
 Art Unit 1642
 REM-3A18 (Room)
 REM-3C18 (Mail Box)

RECEIVED
 DEC 27 2005
 LACH/CHEN, DIVISION
 (STIC)

 Searcher: _____
 Searcher Phone: _____
 Date Searcher Picked up: _____
 Date completed: _____
 Searcher Prep Time: _____
 Online Time: _____

 Type of Search
 NA# _____ AA#: _____
 S/L: _____ Oligomer: _____
 Encode/Transl: _____
 Structure #: _____ Text: _____
 Inventor: _____ Litigation: _____

 Vendors and cost where applicable
 STN: _____
 DIALOG: _____
 QUESTEL/ORBIT: _____
 LEXIS/NEXIS: _____
 SEQUENCE SYSTEM: _____
 WWW/Internet: _____
 Other (Specify): _____

THIS PAGE IS BLANK